



**SAPIENZA**  
UNIVERSITÀ DI ROMA

**Università degli Studi di Roma "La Sapienza"**  
**Scuola Speciale per Archivisti e Bibliotecari**  
**Corso di Diploma per Bibliotecari**  
Insegnamento di Scienza della biblioteca e dell'informazione

## **Folksonomies**

### **Indicizzazione e classificazione delle risorse in rete**

**Candidato**  
**Dott.ssa Barbara Lauro**  
**n° matricola 632480**

**Relatore**  
**Prof.ssa Maria Teresa Biagetti**

**Correlatore**  
**Prof. Giovanni Solimine**

**Anno Accademico 2008/2009**

a Paolo

## Indice

<a href="#">INTRODUZIONE</a>	<a href="#">4</a>
<a href="#">LA DIMENSIONE COLLABORATIVA DEL WEB</a>	<a href="#">4</a>
<a href="#">CAP. I - I TAG</a>	<a href="#">10</a>
<a href="#">CAP. II - LE FOLKSONOMIES</a>	<a href="#">26</a>
<a href="#">CAP. III - FOLKSONOMY vs TASSONOMIE</a>	<a href="#">53</a>
<a href="#">CAP. IV - ANALISI DI UN ESEMPIO: TRECCANI.IT</a>	<a href="#">77</a>
<a href="#">CONCLUSIONI</a>	
<a href="#">    104</a>	
<a href="#">BIBLIOGRAFIA</a>	
<a href="#">    107</a>	

## **INTRODUZIONE**

### **LA DIMENSIONE COLLABORATIVA DEL WEB**

Nel dicembre del 2006 il prestigioso magazine americano *Time* ha eletto il popolo della rete come "Personaggio dell'anno".

Per comprendere a pieno le motivazioni alla base di questo prestigioso riconoscimento è necessario fare un passo indietro e ragionare sui cambiamenti che hanno caratterizzato le dinamiche della rete negli ultimi anni; fin dalla sua nascita, infatti, il processo evolutivo che ha caratterizzato il web si è articolato in modo tanto rapido quanto imprevedibile, ma l'impianto generale di Internet è sempre rimasto stabile e consolidato: da una parte i produttori di contenuti e dall'altra gli utenti, semplici fruitori delle risorse messe a disposizione sul web.

Negli ultimi anni, però, questa storica divisione è divenuta sempre meno percepibile. Il classico paradigma produttore-consumatore è stato progressivamente soppiantato da una visione nuova della Rete, più democratica, in cui anche ad utenti anonimi viene data la possibilità di contribuire direttamente all'evoluzione dei contenuti presenti in Rete attraverso materiale generalmente autoprodotta. Questo è il cambiamento percepito e sottolineato da Time: l'utente si pone al centro del Web, produttore dei contenuti e vera spina dorsale di piattaforme come Wikipedia, Flickr, YouTube, del.icio.us, applicazioni profondamente eterogenee ma caratterizzate da un comune denominatore, la collaboratività.

Condividere informazioni e collaborare per un obiettivo comune: è questo uno dei principi cardine del Web 2.0 e che si rispecchia in tutte le attività che lo caratterizzano, a partire dalla pubblicazione di articoli su Wikipedia,

passando per la pubblicazione di foto, video personali e podcast, fino ad arrivare alla annotazione di contenuti su del.icio.us.

Proprio l'annotazione delle risorse della rete, generalmente indicata col termine di tagging, è una delle tendenze più recenti del Web: l'utente fruisce di contenuti ed individua delle parole-chiave, dette appunto *tag*, che ne denotino al meglio l'informazione convogliata. Un meccanismo tanto immediato quanto utile poiché permette di ovviare ai due limiti principali del Web: da un lato l'assenza di una semantica formalmente definita che descriva il contenuto dell'informazione presente in Rete (pensiamo ad esempio alla poca precisione delle ricerche su termini ambigui, es: pesca, mosca, ecc.), dall'altro la crescita esponenziale ed incontrollata dei contenuti presenti su Internet, inquadrata col termine di *Information Overload*, sovraccarico cognitivo.

L'utilizzo dei tag per annotare le risorse, organizzarle ed associare loro una semantica che ne denoti il contenuto convogliato non è l'unico e probabilmente neppure il più efficace tra quelli attualmente esistenti per eliminare le problematiche legate all'Information Overload, ma si tratta senza dubbio di un approccio innovativo, che si presta ad essere analizzato da vari punti di vista, non ultimo quello della Scienza dell'informazione e della Biblioteconomia.

Il contenuto di documenti digitali viene normalmente indicizzato tramite parole chiavi, e tradizionalmente questa indicizzazione viene fatta da professionisti per facilitare l'accesso e l'organizzazione delle informazioni. Ma quando si ha a che fare con un numero enorme di risorse quali quelle rese disponibili sui siti internet, che hanno cominciato a generare una straordinaria quantità di contenuti attraverso wiki e altri strumenti sociali, è molto difficile per dimensioni, costo e tempo, tenere il passo

nell'indicizzazione.

Quindi l'utilizzo di classificazioni generate dagli utenti, dette *folksonomies*, consente agli utenti di organizzare il proprio contenuto digitale, per trovare gli oggetti di interesse, conservare e mantenere il catalogo utilizzando i cosiddetti *tag*, etichette associate alle risorse per indicarne il contenuto.

Il termine *folksonomy* è un neologismo di origine inglese coniato nel 2004 da Thomas Vander Wal, dall'unione di "taxonomy" (tassonomia), la scienza che si occupa dei modi di classificazione di qualsiasi concetto, cosa, o essere vivente, seguendo un ordine gerarchico (dal greco: taxis = ordine, nomos = regole) e "folk" (gente)

Il termine è stato quindi utilizzato per indicare in genere una classificazione spontanea e collaborativa, creata dagli utenti del web che, assegnando dei *tag* (etichette), classificano immagini, video, link, testi e così via.

Si tratta dunque di una forma di tassonomia "inversa" creata da chi la usa (la gente, "folk") che, spontaneamente e liberamente, organizza e "etichetta" le informazioni (pagine web, foto, video, etc.) rese disponibili attraverso la rete.

Questo sistema di "etichettatura" è stato in genere identificato come una forma di classificazione e il termine *folksonomy*, denominata anche "social tagging", "collaborative tagging", "social indexing", tassonomia popolare, etnoclassificazione, etc. a seconda all'aspetto sociale, informativo o tecnologico del processo che si vuole sottolineare, si è rapidamente affermato in rete come un sistema di annotazione collaborativa di informazioni mediante l'utilizzo di tag scelti liberamente dagli individui che ne fruiscono.

Un aspetto però su cui non si è riflettuto abbastanza in merito a questo

fenomeno è se si tratti effettivamente di una forma di classificazione o piuttosto di un sistema di indicizzazione delle risorse, dato che l'attività di associazione di tag concerne più il concetto di "aboutness" che non quello del dominio di conoscenze.

Il fatto che fra i termini utilizzati per definire questo fenomeno si sia affermato proprio il neologismo *folksonomy* (tradotto in italiano come "classificazione popolare") ha infatti portato i non addetti ai lavori a parlare inconsapevolmente di "tassonomia" e di "classificazione", utilizzando una terminologia che appartiene più al mondo dell'archivistica che non a quello delle biblioteche. Volendo invece indagare il fenomeno in termini strettamente biblioteconomici, si rileva come la pratica di attribuzione dei tag alle risorse in rete sia più vicina ad una forma di indicizzazione, e più precisamente ad una indicizzazione per soggetto, che non ad una classificazione.

Si tratta in ogni caso di una indicizzazione basata su termini descrittivi del tutto priva di struttura e di relazioni necessarie tra gli elementi. È una catalogazione spontanea e collaborativa di risorse eterogenee che si adatta con facilità ad ambienti di collaborazione libera com'è la comunità del web.

Lanciati su vasta scala all'inizio del 2005 da Technorati <<http://technorati.com>>, il principale motore di ricerca dedicato al mondo dei blog, i tag sono stati sorprendentemente adottati da un gran numero di siti personali, ancora prima della diffusione di sistemi automatici per la loro gestione, ma anche di applicazioni internet basate sulle reti sociali quali Flickr o Youtube, dando avvio a una delle più sorprendenti propagazioni di "buone pratiche" che la rete abbia conosciuto negli ultimi tempi.

Le caratteristiche della folksonomy suggeriscono importanti riflessioni di ordine sia metodologico che tecnico, oltre a stimolare interessanti intuizioni sulle dinamiche sociali che prendono vita nel web 2.0, dove il *tagging* (l'applicazione di "etichette" ai contenuti) è "social", poiché viene abitualmente affidato agli utenti. Considerato che gli organizzatori dell'informazione sono di solito gli utenti finali, la folksonomia dovrebbe produrre risultati che riflettono in maniera più definita l'informazione secondo il modello concettuale della popolazione che la utilizza.

Sono in molti a pensare che, posto un universo di dati potenzialmente infinito, quale è oggi la Rete, il miglior disordine possibile consiste nella buona pratica di "associare delle etichette" alle risorse condivise sul web, siano esse testi, immagini, indirizzi web o video.

Afferma Sergio Maistrello nel descrivere il fenomeno delle folksonomies nel suo libro *La parte abitata della rete*:

"Il ferreo regime di catalogazione che regna tra gli scaffali di una biblioteca è ordine; la tavola periodica degli elementi chimici è ordine; la classificazione biologica degli esseri viventi è ordine. Sono tutti esempi di un ordine oggettivo pre-codificato, che richiede la condivisione di criteri di interpretazione e il rispetto di precise sequenze gerarchiche, che si imparano a riconoscere per lo più in seguito a studi specialistici. La folksonomy, invece, prescindendo da qualunque schema preordinato, è disordine dotato di una buona segnaletica."<sup>1</sup>

La folksonomy però non è perfetta e viene costantemente migliorata nella pratica e negli strumenti. Il limite più evidente è dato dall'ambiguità delle

---

<sup>1</sup> Sergio Maistrello, *La parte abitata della Rete*, Milano, Tecniche Nuove, 2007, pag. 84



catalogazioni spontanee (persone differenti classificano in modo differente le stesse risorse) e dall'uso di sinonimi, di luoghi comuni passeggeri, di nomi che possono essere scritti in modi diversi pur essendo riferiti allo stesso concetto. Quella che dal punto di vista della classificazione può sembrare causa di entropia in un ambiente per sua natura già caotico, in realtà non fa altro che riprodurre dentro la Rete una struttura di organizzazione per analogie molto simile a quella che sta alla base del ragionamento umano.

Le tecniche per ovviare a questi tipi di problematiche vanno studiate a fondo per rendere più efficace i sistemi di tagging condiviso. Sul portale dell'IstitutoTreccani <[www.treccani.it](http://www.treccani.it)>, ad esempio, si è scelto di proporre agli utenti un set di termini fra cui scegliere quelli da attribuire alle singole risorse enciclopediche, parole che rappresentano a loro volta altre risorse, fino ad ottenere una rete di correlazioni fra termini costruita dai navigatori.

Se comunque accettiamo una visione della rete secondo conversazioni generate da individui, cosa che già avviene nel web 2.0, la conquista a cui porta la folksonomy prescinde da tutti i suoi vantaggi e svantaggi contingenti. Semplicemente bisogna smettere di ricercare un ordine dentro il web e iniziare ad accettare l'idea che la rete abbia bisogno di disordine per funzionare al meglio.

## **CAP. I - I TAG**

### **1.1. Cosa sono e a cosa servono**

Il termine inglese *tag* ( propr. “coda”, ma anche “cartellino, contrassegno”), è usato per indicare diversi concetti. Nel linguaggio jazzistico, ad esempio, indica una frase musicale (detta anche, con termine francese, *queue*, cioè coda) che si aggiunge al ritornello, ma può essere usato anche per indicare la sigla o la firma stilizzata che identifica l'autore di un graffito come vero e proprio contrassegno di identità.

In informatica invece il tag è l'unità elementare con cui si marcano i campi di un file (e detta perciò anche marcatore) al fine di poterli identificare e successivamente processare. Nei linguaggi HTML e XML esistono tag che identificano singoli elementi della pagina come titolo, autore, un link, etc.

I tag sono quindi dei metadati, dei dati che spiegano altri dati.

La logica di utilizzare “etichette” per identificare porzioni di informazioni si è rapidamente affermata sul web anche al di fuori dei linguaggi di programmazione fino ad utilizzare questo termine per intendere in senso più generale una parola associata ad una risorsa della rete (un'immagine, una mappa geografica, un post, un video, etc), che descrive l'oggetto rendendo possibile la classificazione e la ricerca di informazioni basata su parole chiave.

Bettina Berendt e Christoph Hanser in un articolo dal titolo “Tags are not metadata, but “just more content” - to some people”<sup>2</sup>, disegnano una

---

<sup>2</sup> Berendt, Bettina e Hanser, Christoph, *Tags are not metadata, but “just more content” - to some people*, International Conference on Weblogs and Social Media, Boulder, Colorado, USA, 2007  
<http://www.icwsm.org/papers/2--Berendt-Hanser.pdf>

particolare tassonomia dei tag, una classificazione gerarchica delle differenti tipologie di etichette che l'utente può utilizzare per annotare le informazioni della Rete. Secondo questo studio, i tag possono essere suddivisi in cinque diverse categorie:

1. Content-Based Tag – si tratta dei tag che indicano il contenuto dell'oggetto o delle categorie cui l'oggetto appartiene (es. Auto, Ferrari, ecc.)
2. Context-Based Tags – forniscono informazioni legate al contesto dell'annotazione (es. vacanze Natale, Agosto, ecc.)
3. Attribute Tags – caratteristiche dell'oggetto che non sono direttamente derivate dalla categoria cui appartiene (es. la mia macchina)
4. Subjective Tags – tag soggettivi, che esprimono impressioni personali ed emozioni sull'annotazione effettuata (es. bello, interessante)
5. Organizational Tags - tag che esprimono annotazioni personali, slegate però dall'aspetto emotivo (es. da leggere)

I tag, che in genere vengono scelti direttamente dagli autori/creatori dell'oggetto dell'indicizzazione in base a criteri del tutto informali, sono stati presto associati al concetto di web 2.0 e ai cosiddetti servizi di social bookmarking.

I tag sono uno strumento del web 2.0 perché consentono a chi pubblica una risorsa di etichettarla ("taggarlo" in *slang* 2.0) con una o più parole che ne definiscano e ne facciano capire il contenuto ed il senso.

Sono qualcosa di più quindi delle parole chiave, *keywords*, perché non vengono estrapolati automaticamente dal testo, dal titolo o dalla descrizione del contenuto ma assegnate direttamente o da colui che l'ha pubblicato e che per primo è in grado di descriverne il significato o da chi fruisce di quel contenuto.

L'uso dei tag infatti è molto intuitivo: ogni unità di contenuto immessa in circolo può essere integrata con alcune meta-informazioni semantiche che ne specificano i temi impliciti ed espliciti, il contesto e tutto quanto può essere utile per archivarla nella grande rete e favorirne il reperimento da parte delle persone interessate.

Se nel caso delle parole chiave siamo di fronte ad una ricerca automatizzata, qui c'è un prezioso intervento umano che assegna un senso ad un contenuto. Il vantaggio per chi cerca qualcosa attraverso i tag è di avere la possibilità di provare ad effettuare ricerche semantiche formate quindi anche da insiemi di parole e basate sul senso della frase, non solo sul significante ma anche sul significato della stessa.

Molti studiosi considerano i tag come veri e propri metadati, perché descrivono, classificano, organizzano altri dati.

Un metadato infatti è un tipo di informazione che descrive un insieme di dati, necessario per ricercare i dati, rappresentarli e utile a comprenderli.

Bettina Berendt e Christoph Hanser, invece, nel loro già citato contributo alla International Conference on Weblogs and Social Media scrivono:

In a corpus of posts consisting of body elements (text, title, ...) and author tags, the tags are not metadata but content if the tags have a low similarity with the body (such that body features cannot be used to predict the tags, or vice versa), and the combination of body and tags predicts the human consensus classification<sup>3</sup>.

Nell'esaminare il valore aggiunto fornito dai tag rispetto ai contenuti a cui vengono associati è bene però considerare il fatto che queste etichette

---

<sup>3</sup> *ibidem*

possono essere di due tipi: oggettivi e soggettivi.

- tag soggettivo - il significato ha senso solo per colui il quale assegna il tag (per esempio foto di Bali durante il viaggio di nozze con tag “honeymoon”);
- tag oggettivo - il significato ha senso per colui che assegna il tag, ma è condivisibile da tutti, e conseguentemente all’oggetto viene attribuito lo stesso tag da una moltitudine di persone (per esempio tag “pesce” per la foto di un pesce).

Fra le due tipologie quella dei tag oggettivi, per il fatto che vengono ripetuti e usati da più utenti, crea un’informazione con un elevato valore democratico di riconoscimento di una qualità, un attributo generalmente condiviso e condivisibile. La corrispondenza tra tag e qualità della risorsa taggata è quindi molto probabile che venga considerata corretta anche da chi, facendo ricerche, troverà quel soggetto come risultato di una ricerca per quella determinata parola chiave.

Quindi, dalla somma delle classificazioni personali fatte per scopi individuali si ottiene una classificazione collettiva al servizio della comunità.

In più l’utilizzo dello stesso tag oggettivo condivisibile da parte dell’ennesimo visualizzatore alimenterà l’affermazione democratica dello stesso tag tra tutti gli utenti, generando e alimentando nel tempo un circolo virtuoso/vizioso (questione di punti di vista...) che spinge verso la nascita di gerghi e slogan (per esempio lo stesso “Web 2.0”).

Il sistema di tagging è poi in genere accompagnato da una serie di altre funzionalità per favorire interazioni cooperative come commenti, raccomandazioni di utenti con analoghi interessi, tag correlati, possibilità di costruire gruppi omogenei.

## 1.2. **Come sono nati**

I tag intesi come “etichette” associate ai contenuti sono nati da un esperimento condotto presso la Carnegie Mellon University su un software sviluppato per testare le capacità extrasensoriali delle persone, detto ESP Game, allo scopo di migliorare la ricerca di immagini su Internet.

L'assegnazione di tag da parte degli utenti è stato infatti proposto come modo per migliorare la qualità della ricerca di immagini sul web da Von Ahn e Dabbish, coloro che hanno introdotto l'ESP Game <[www.espgame.org](http://www.espgame.org)> un "gioco" per coinvolgere gli utenti e spingerli a fare il lavoro di migliaia di operatori: fornire "tag" accurati per aumentare la precisione della ricerca immagini.

Il programma offriva la possibilità a due diversi utenti di sfidarsi nell'associare descrizioni alle immagini. Durante le sessioni di gioco, della durata di 90 secondi, ai due utenti veniva infatti proposto lo stesso set di foto e ciascuno dei due aveva la possibilità di guadagnare punti soltanto quando la sua descrizione coincideva con quella dell'altro concorrente. In questo modo vincono le coppie che sono più "in sintonia", con classifiche costantemente aggiornate.

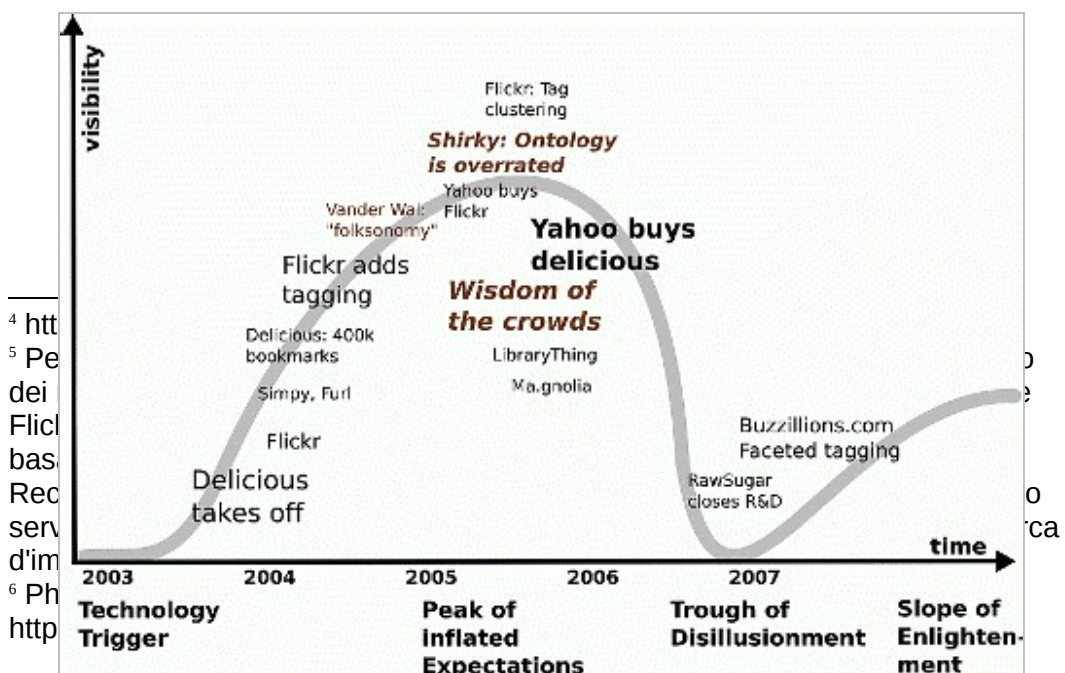
In questo modo, i tecnici di ESP Game sono riusciti a catturare dati importantissimi per elaborare la rilevanza statistica di ciascun tag, al fine di ottimizzare la ricerca ed ottenere un database categorizzato d'immagini digitali.

Il software ESP Game sviluppato alla Carnegie Mellon University è stato poi utilizzato da Google, sempre sotto forma di gioco, il Google Image

Labeler<sup>4</sup>, per assegnare descrizioni contestuali per ciascun file del database fotografico di Google Image Search.<sup>5</sup>

Una forma di tagging su più larga scala è stato poi lanciato da Google nella sua piattaforma di webmail, Gmail, che permette di assegnare ai messaggi una o più etichette, sostituendo a una struttura gerarchica in “cartelle”, tipica del file system di un PC, un’organizzazione orizzontale e multicategoriale.

La diffusione dell’uso dei tag ha poi seguito una crescita esponenziale a partire dal 2003, con il lancio di delicious e Flickr, che ha segnato il momento di slancio e di grande visibilità del “social tagging”, come descritto nell’immagine sulla storia dell’uso dei tag di Philipp Keller, ricercatore presso Instituto de Investigaciones Filosóficas dell’Università del Messico, in un suo articolo<sup>6</sup> sulla storia del tagging, presentando un diagramma e una cronologia degli avvenimenti che hanno coinvolto il tagging.



In ambito scientifico, invece, uno dei primi studiosi di Scienze dell'informazione ad essersi occupato dei sistemi di "tagging" è stato Adam Mathes in un articolo del 2004<sup>7</sup> scritto per il corso di Scienze dell'informazione e della biblioteca della University of Illinois Urbana-Champaign, poi ampiamente elogiato e citato in tutti gli studi sulle folksonomy.

Nel suo articolo sulla "classificazione cooperativa e comunicazione attraverso metadata condivisi" Mathes fornisce la prima indagine sul "social tagging" e sui sistemi di folksonomy e descrive la funzionalità di questi sistemi, confrontando il sistema di etichettatura in Flickr e del.icio.us. Mathes mette quindi in evidenza i limiti del sistema di tagging da parte degli utenti che riguardano in particolare "an uncontrolled vocabulary, ambiguity, inconsistent orthography, and synonymy".

Mathes afferma inoltre che le folksonomies funzionano grazie ad elementi di partecipazione e condivisione propri delle reti partecipative del web:

It is difficult to define a metric by which one could argue folksonomies are a success or failure, but the degree that it does seem to be effective in these systems as a way of organizing information, and that a large group of people are using these systems, I posit, is due to a few important factors. The overall costs for users of the system in terms of time and effort are far lower than systems that rely on complex hierarchal classification and categorization schemes. In addition to this structural difference, the context of the use in these systems is not just one of personal organization, but of communication and sharing. The near instant feedback in these systems leads to a communicative nature of tag

---

<sup>7</sup> Mathes, Adam, *Folksonomies - Cooperative Classification and Communication through Shared Metadata*, 2004  
<http://www.adammathes.com/academic/computer-mediated-communication/folksonomies.html>



use.

I punti di debolezza delle folksonomy possono dunque rappresentare, allo stesso tempo, dei punti di forza, quale il costante aggiornamento, merito della velocità e semplicità con cui gli utenti “taggano” le risorse, la possibilità di inserire una stessa risorsa in categorie multiple, contrariamente a quanto avviene nei sistemi di categorizzazione tradizionali e , infine, la possibilità di vedere come gli utenti “taggano” le risorse, cioè lo schema di categorizzazione usato e tracciare con rapidità l'emergenza di alcuni temi rispetto ad altri.

### 1.3. **Chi li utilizza**

La pratica di assegnare un tag alle risorse ha trovato naturale applicazione nella gestione e recupero di oggetti non testuali, come le fotografie, le immagini, i video, la musica o gli stessi siti web, con la pratica del cosiddetto “social bookmarking”, un servizio che rende disponibili elenchi di segnalibri (bookmark) creati dagli utenti.

I siti di social bookmarking sono stati fra i primi ad organizzare il loro contenuto tramite l'uso di tag. La popolarità di questi siti è in costante crescita, in quanto sono uno strumento facile e intuitivo per individuare, classificare, ordinare e condividere le risorse in rete attraverso la pratica dell'etichettatura. Ulteriori risultati possono inoltre essere ottenuti raggruppando insieme più tag.

In un sistema di social bookmarking gli utenti registrano un catalogo di risorse Internet ritenute, a vario titolo, utili e stimolanti. Sovente questi elenchi sono pubblicamente accessibili e altre persone con analoghi interessi possono visionarli, casualmente o suddivisi per categorie.

La categorizzazione delle risorse avviene per mezzo di tag liberamente e informalmente scelti dall'utente. La maggioranza dei servizi di social bookmarking permette agli utenti la ricerca di segnalibri associati a determinate etichette, oltre all'ordinamento delle singole risorse Internet in base al numero delle segnalazioni. Molti siti hanno inoltre implementato algoritmi per ottenere risultati da ricerche condotte su gruppi di parole chiave.

Il concetto di "condivisione in rete dei bookmark" è databile verso la metà del 1999, ma le prime proposte in questo senso furono di breve durata, travolte dalla crisi che investì il comparto dell'alta tecnologia dopo il 2001. Il social bookmarking diviene una pratica diffusa solo nel settembre del 2003 con la nascita del sito del.icio.us <<http://del.icio.us>>, che ha fatto da apripista e si qualifica ancora oggi come punto di riferimento.

Del.icio.us è stato infatti il primo servizio di bookmarking che ha permesso di creare segnalibri di Url, come i "preferiti" nel proprio browser, ma caricandoli in un ambiente condiviso dove tutti possono vederli e di associarvi termini significativi in modo da poterli ricercare per argomento. I dati sono qui rappresentati non solo da pagine di siti web ma anche da documenti testuali digitali tradizionali, quali articoli, libri, report, oltre a news, posting dei blog ecc.

In Delicious quindi gli utenti salvano i propri url e possono taggarli con una loro descrizione per ritrovare i loro indirizzi preferiti. Il sito mostra anche come gli altri utenti abbiano taggato il materiale creando in questo modo una classificazione data dalla comunità e non più dal singolo.

Come funziona? Dopo la registrazione è possibile iniziare il bookmarking dei propri indirizzi compilando un form e inserendo URL, titolo, descrizione e i tag.

Una possibilità interessante che offre Delicious è la sottoscrizione ai tag, attraverso la quale si può ricevere una notifica ogni volta che qualcuno salva una pagina con un determinato tag.

Se ad esempio si vuole seguire un tag di particolare interesse, ci si "iscrive" a quel tag e con l'iscrizione si decide di "filtrare" i link di del.icio.us in modo da leggere solo quelli degli argomenti preferiti. È inoltre possibile raggruppare tag simili in cartelle, utilizzando le opzioni per l'ordinamento

dei tag.

La stessa cosa è stata realizzata per le fotografie dal popolare sito di immagini Flickr <<http://www.flickr.com>>, acquisito da Yahoo! nel 2005.

Flickr permette alle persone di condividere le proprie foto su internet. Altre persone possono commentare le immagini pubblicate e assegnare tag alle foto. Per utilizzare Flickr si richiede anche agli utenti di creare un account. C'è anche l'opzione di pagare un abbonamento per usufruire di un account con caratteristiche avanzate, come maggiore spazio per le fotografie.

In Flickr, a differenza di Delicious, gli utenti taggano le proprie immagini per un uso personale e i tag permettono all'utente di navigare all'interno del proprio catalogo tramite i termini utilizzati per la categorizzazione.

Flickr non permette tag multipli e per unire due parole in un tag si usano le virgolette doppie, il tag viene applicato nell'upload della foto. Su Flickr coesistono diversi formati di contenuto (immagini, video, mappe).

Flickr contiene un insieme di metadata, che permettono alle immagini di essere categorizzate in base, ad esempio, a chi l'ha scattata, o alla sua posizione geografica, in modo tale che la navigazione diventa qualcosa di non prevedibile, però può diventare un elemento sociale, nel momento in cui più utenti collaborano.

Per i video, invece, si è affermato Youtube <<http://www.youtube.com>>.

Su Del.icio.us, Flickr o Youtube ciascuno può caricare i propri link preferiti, o le fotografie, o i video, classificandoli con tag e con descrizioni per condividerli con gli amici o con qualunque altro utente sconosciuto.

Su un versante più serio, ma con gli stessi identici meccanismi, il gruppo editoriale della prestigiosa rivista «Nature» ha lanciato Connotea

<<http://www.connotea.org>>, un servizio progettato per i ricercatori e per i bibliotecari, per il tagging di articoli, pagine Web e altro materiale, di ricerca o di studio.

Connotea è un servizio particolare di social bookmarking pensato appositamente per la comunità scientifica internazionale. Unisce i servizi di tagging a un target scientifico e servizi innovativi quali commentare gli URI inseriti con possibilità di esportare automaticamente la propria library in formato RIS (usato per le references nei paper scientifici).

Alla fine del 2005 Connotea è uscito dalla fase di beta test e ora raccoglie più di 30 mila link, annotati con 17 mila tag, su diversi argomenti, dalla medicina all'informatica.

Grazie all'uso dei tag in Connotea è possibile

- consultare la lista dei tag utilizzati da ogni utente
- editare dei tag
- ricevere suggerimenti di tag correlati
- sapere quali utenti hanno archiviato le stesse risorse o hanno usato gli stessi tag.

Il sistema è stato studiato anche per utilizzare i DOI, i Digital Object Identifier (codici univoci per catalogare le risorse Internet), inserendo i quali Connotea riporta automaticamente le esatte citazioni bibliografiche, con collegamenti alla grande libreria on line Amazon, a riviste come «Nature» e «Science», o all'archivio medico statunitense Pubmed.

Connotea, come del resto Del.icio.us, prevede anche il geotagging, per legare gli articoli alle coordinate di latitudine e longitudine, per esempio se si tratta di materiale che parla di un luogo ben preciso; fatto questo si può utilizzare il programma Google Earth per visualizzarlo sulle sue mappe<sup>8</sup>.

---

<sup>8</sup> Connotea collabora con un altro servizio di questo tipo, sempre dedicato ai bibliotecari e agli accademici, Citeulike <<http://www.citeulike.org>>.

I due sistemi stanno mettendo a punto degli standard comuni, in modo che i loro

I software di tagging sono spesso dotati di servizi aggiuntivi, come quelli offerti da Connotea, che permettono di costruire citazioni, cioè descrizioni, più o meno complesse, dei documenti salvati, recuperando quando possibile i dati automaticamente (Connotea, per esempio, può interpretare il DOI), di visualizzare i tag assegnati alla risorsa da altri utenti e i tag correlati, di scorrere le collezioni di siti archiviati da altri utenti che si occupano di argomenti simili, di costituire gruppi di utenti con presunti interessi comuni e, tramite RSS, di monitorare i cambiamenti nelle aree di interesse.

Anche nelle modalità di assegnazione dei tag vi sono differenze importanti. Del.icio.us permette di assegnare solo tag composti da una sola parola, mentre Connotea applica la tecnica della frase delimitata da virgolette, tipica dei motori di ricerca. Al momento dell'assegnazione dei tag, il primo fa vedere i tag più popolari assegnati da altri all'indirizzo web che si vuole salvare, mentre Connotea ripropone solo quelli utilizzati dall'utente.

Ognuno di questi servizi di bookmarking, quindi, ha sviluppato in questi anni un suo sistema di tagging, arrivando a costruire il sistema che più si addice alle esigenze del suo target di utenza e agli obiettivi che tramite i tag si vogliono ottenere nella organizzazione le risorse scientifiche in rete.

#### **1.4. Quasi un ossimoro: i tag “ufficiali”**

Un tag ufficiale è una parola chiave adottata nell'ambito di eventi o conferenze in modo che i partecipanti associando a propri documenti,

---

archivi siano compatibili.

video, immagini questo tag rendono il proprio materiale facilmente individuabile dai motori di ricerca e, pertanto, facilmente condivisibile con gli altri partecipanti all'evento.

Per promuovere in Rete qualsiasi tipo di evento è ormai inevitabile confrontarsi con le possibilità aperte dalle forme di produzione diretta di contenuti da parte del pubblico e dalla condivisione in tempo reale che l'adesione ai social network rende possibile.

Se però non si può prescindere dall'utilizzare il contributo dei navigatori su piattaforme come YouTube, Flickr o Facebook, quando si vuole promuovere un evento a cui gli utenti partecipano in prima persona è importante "guidare" in qualche modo la produzione di questi contenuti in modo da rendere possibile la loro associazione all'evento stesso. Partecipare ad un concerto, ad esempio, significa non solo essere presente nel luogo e nel momento dello spettacolo, ma produrre comunicazione su questo e renderla visibile e, soprattutto, raccordata ad altra comunicazione prodotta sullo stesso fatto.

Se ciò che rende possibile questo raccordo è la logica del tag, diventa sempre più importante ragionare su una strategia di coordinamento sulle modalità di tagging, e in molti hanno fatto ricorso all'uso di quelli che sono generalmente definiti come "tag ufficiali" di un evento. Si tratta di parole chiave adottate spesso nell'ambito di eventi o conferenze in modo che i partecipanti associando a propri documenti, video, immagini questo tag rendono il proprio materiale facilmente individuabile dai motori di ricerca e, pertanto, facilmente condivisibile con gli altri partecipanti all'evento.

Parlare di "ufficialità" di tag potrebbe sembrare una contraddizione, dato che queste "etichette" del web 2.0 sono sempre state considerate come contributo "libero" da parte degli utenti alla etichettatura delle informazioni

presenti sul Web. Eppure si stanno diffondendo in maniera sempre più evidente dei sistemi di etichettatura che usano i tag proprio per dare ordine al disordine tipico dei sistemi Content User Generated.

Prendiamo ad esempio il gruppo rock statunitense dei R.E.M. che, seguendo le logiche pure delle culture partecipative, in un'ottica di convergenza fra produzione e consumo, hanno stimolato i propri fan a condividere i contenuti prodotti dal basso durante il loro tour mondiale.

Durante l'ultimo concerto di Torino, ad esempio, sui maxischermi dietro il palco sono comparse le scritte: Flickr & YouTube tag: remtorino e Blog & Twitter tag: #remtorino con cui la band ha invitato i fan a utilizzare l'attributo "remtorino" per etichettare i video, le immagini, i blog, i messaggi Twitter pubblicati sulle varie piattaforme UGC.

Un invito che ritorna anche sul sito ufficiale dei R.E.M., dove è stata aperta una sezione ad hoc dedicata a raccogliere i contenuti multimediali prodotti dai fan nelle varie date del tour e viene spiegato come taggare i video prodotti con i cellulari, le foto fatte e i messaggi scambiati via twitter:

For example, for R.E.M.'s show in Berlin, Germany, you would do the following:

In Flickr, tag your photos "rememberlin"

In Twitter, add "#rememberlin" to your Tweets

On YouTube, tag your video "rememberlin"

If you blog about the show, please add "#rememberlin" to your blog entry.

Si tratta evidentemente di una intelligente apertura al cambiamento in atto relativo all'intreccio tra "essere fan" e UGC, ma anche di un uso efficace della logica delle folksonomy nell'aggregazione e classificazione di contenuti generati e diffusi in innumerevoli diversi punti della rete Web.

Altri esempi di uso di tag ufficiali si possono ritrovare nel modo di gestire la



musica introdotto da last.fm <<http://www.last.fm>> dove esistono sia *tag ufficiali* che *tag personali* che permettono di gestire playlist miste, o in un sistema come Moodle <<http://moodle.org>> che gestisce due modalità di tagging, una sociale e trasversale al sistema e una contestuale a sottoinsiemi (corsi di formazione).

## **CAP. II - LE FOLKSONOMIES**

### **2.1. Il ruolo dell'utente nel web 2.0**

Il concetto di web 2.0 è nato durante una sessione di brainstorming all'interno di una conferenza tra O'Reilly Media, casa editrice americana specializzata in pubblicazioni riguardanti le nuove tecnologie e Internet in particolare, e MediaLive International. In quell'occasione Dale Dougherty (dello staff O'Reilly), riflettendo sulla progressiva importanza del web nella vita quotidiana e sul diffondersi di nuovi servizi di alta qualità, coniò il termine che fu subito accettato dal gruppo. Nacque così la "Web2.0 Conference" rinominata in seguito "Web2.0 Summit", che dall'ottobre 2004 è il luogo in cui vengono analizzati e censiti lo sviluppo del web, i trend e le nuove possibilità di business del settore.

La definizione di web 2.0 che si è poi affermata in rete e che oggi è più largamente accettata è quella di Tim O'Reilly, che in un lungo articolo presentato sul suo blog<sup>9</sup>, traccia le linee guida di questo fenomeno, indicandone caratteristiche innovative, peculiarità, e fornendone una definizione piuttosto compatta:

"Web 2.0 is the business revolution in the computer industry caused by the move to the Internet as platform, and an attempt to understand the rules for success on that new platform".

Da questa definizione si può capire il fulcro del web 2.0: vedere la rete come una "piattaforma", sfruttandola il più possibile. Con questa

---

<sup>9</sup> Tim O'Reilly, *What is Web 2.0*, articolo originale tratto dal blog <http://www.oreillyn.com/pub/a/oreilly/tim/news/2005/09/30/what-is-web-20.html>

espressione viene inteso il progressivo (e sempre più intensivo) uso di applicazioni web finalizzato a portare verso una continuativa e sempre maggiore presenza online degli utenti.

All'articolo di Tim O'Reilly sono seguiti dibattiti e diatribe in rete su cosa sia realmente il web 2.0.

Si tratta senza dubbio di dibattiti fisiologici alla luce dell'oggettiva difficoltà a definire con precisione un qualcosa che, come il web 2.0, è ancora in divenire. Ai fini della nostra ricerca è però più utile fornire una definizione di web 2.0 maggiormente "operativa", che permetta di apprezzare le caratteristiche distintive dei servizi 2.0, comprendere quali funzionalità possano assumere questa dicitura e quali no, e, soprattutto, capire cambiamenti, vantaggi, e, in generale, implicazioni, che il web 2.0 può portare dal punto di vista dell'utente.

"Web 2.0 is a set of economic, social and technology trends that collectively form the basis for next generation of the Internet – a more mature, distinctive medium characterized by user participation, openness, and network effects"<sup>10</sup>

Partecipazione degli utenti, apertura ed effetti di rete: queste le caratteristiche distintive del Web 2.0, sempre secondo Tim O'Reilly. Il web 2.0 esprime dunque tecnologie in grado di incrementare le possibilità di partecipazione ed approfondimento, dando agli utenti una dimensione e soprattutto un ruolo nuovo, più attivo e determinante.

Questa definizione apre la strada ad una serie di riflessioni relative al ruolo e al contributo che gli utenti possono assumere nell'ambito delle dinamiche della rete e nell'organizzazione delle sue risorse.

---

<sup>10</sup> *ibidem*

Se infatti il denominatore comune che caratterizza tutti i servizi del web 2.0 è la partecipazione dell'utente, la collaboratività, la possibilità di contribuire all'evoluzione della rete e di mettere a disposizione degli altri ciò che ognuno degli utenti produce, è giusto chiedersi se il contributo dell'utente può migliorare la natura dei servizi offerti dalla rete.

La partecipazione è uno degli elementi essenziali per mettere in pratica la visione del *Read/Write Web* ampiamente trattata in numerosi articoli sul web 2.0, una visione in cui internet deve funzionare in modo duale, con gli utenti che agiscono contemporaneamente da lettori e scrittori ed i Social Software che fungono da veri e propri catalizzatori per mettere in atto questa visione.

Dal punto di vista della tecnologia informatica utilizzata il web 2.0 non rappresenta una rivoluzione, perché ciascuno dei servizi presentati ha un suo corrispettivo perfettamente funzionante nel Web 1.0 o persino sui desktop dei nostri PC. Qual è dunque il fattore determinante alla base del successo di queste applicazioni?

Secondo l'opinione di O'Reilly esiste un denominatore comune che unisce sotto un'unica logica applicazioni così eterogenee e differenti. Come sottolineato in un suo articolo<sup>11</sup>, il fattore che contraddistingue le compagnie web 1.0 sopravvissute alla "bolla" del 2000 e le compagnie web 2.0 più famose è un particolare stile architetturale che caratterizza, indistintamente, le applicazioni più importanti e conosciute, da lui definito "architettura di partecipazione".

L'idea alla base della "architettura di partecipazione" è quella di progettare

---

<sup>11</sup> Tim O'Reilly, *The Architecture of Participation*, articolo tratto dal blog [http://www.oreillynet.com/pub/a/oreilly/tim/articles/architecture\\_of\\_participation.html](http://www.oreillynet.com/pub/a/oreilly/tim/articles/architecture_of_participation.html)

le applicazioni web ponendo particolare attenzione a supportare, all'interno della logica generale, meccanismi finalizzati ad inserire ed integrare contenuti prodotti dagli utenti.

Google, ad esempio, deve il suo successo al PageRank, un algoritmo che è basato sul computo dei link che gli utenti inseriscono all'interno delle proprie pagine. Siti come Wikipedia, Flickr, YouTube e Del.icio.us, senza il contributo degli utenti non avrebbero ragion d'essere.

Quello che dunque il web 2.0 va a stravolgere non è la natura del Web o la tipologia di applicazioni che il web stesso supporta, ma, viceversa, il ruolo totalmente nuovo che l'utente va ad assumere all'interno di questo scenario. Il classico paradigma amministratore/visitatore viene totalmente soppiantato da un modello nuovo, in cui gli utenti delle applicazioni svolgono il ruolo duale di *prosumers*, fusione tra *producers* e *consumers*, produttori di contenuti personali ma allo stesso tempo consumatori di risorse prodotte da altri e messe a disposizione sulla rete.

In questa visione l'utente diviene dunque il vero e proprio valore aggiunto delle applicazioni web 2.0, colui il quale – attraverso semplici meccanismi di votazione o raccomandazione – può determinare il successo e l'insuccesso dei contenuti prodotti da altri, delineando tendenze, mode e novità della rete.

Un'ulteriore analisi delle dinamiche del web potrebbe far emergere la riflessione che il contenuto generato dagli utenti, tutto sommato, non sia qualcosa di così "innovativo". Guardando a ritroso la storia del web, infatti, ci si rende conto come – da sempre - buona parte dei contenuti indicizzati in rete sia opera degli utenti: è la natura stessa del web ad attivare meccanismi di partecipazione attiva e di condivisione delle informazioni.

Perché dunque tutto questo interesse recente verso i contenuti generati

dagli utenti? Non è una novità che l'utente collabori ad estendere il contenuto della Rete attraverso il proprio contributo e che, soprattutto, la maggior parte del contenuto indicizzato sia opera spontanea degli utenti. Perché, come sottolineato da più parti<sup>12</sup>, le principali compagnie hanno aspettato 15 anni prima di rendersi conto che il valore aggiunto in termini di informazione e conoscenza non ce lo mettevano solo i media, ma che, al contrario, la gran parte dei forse 100 miliardi di pagine web era stata creata da volontari, in maniera spontanea?

La recente attenzione verso l'utente "produttore di contenuti" può essere forse motivata in chiave spiccatamente statistica: come mostrato in una analisi del 2006 dell'agenzia Morgan-Stanley, le applicazioni legate al web 2.0 sono quelle che nell'ultimo anno hanno conosciuto la crescita più importante raddoppiando (nel caso di Wikipedia) o persino espandendo in modo esponenziale (come nel caso di YouTube) i contatti registrati nell'arco di un anno solare.

Si tratta di numeri per certi versi quasi sorprendenti, ma che aiutano notevolmente a fornire risposte ragionevoli alla domanda precedente: l'utente produce contenuti per il web da sempre, ma nel momento in cui la mole e la tipologia di queste risorse è divenuta tale da rappresentare buona parte dell'attuale traffico della rete, è naturale che questo dato abbia attirato l'attenzione di tutte le compagnie che investono sul web. Il termine *User-Generated Content*, quindi, nasce proprio con queste due motivazioni: descrivere il fenomeno e denotare tutte le analisi e gli studi finalizzati a comprenderne motivazioni, tendenze, sviluppi futuri ed implicazioni economico-sociali.

---

<sup>12</sup> Chips&Salsa, articoli e appunti da Franco Carlini, *Utilitarista o altruista: il Web Sociale*, <http://chipsandsalsa.wordpress.com/2006/07/06/utilitarista-o-altruista-il-web-sociale/>

L'originaria visione del *Participative Web* di Tim O'Reilly si sta esplicitando, oltre che in differenti contesti, anche sotto varie forme. Gli utenti del web, come detto, possono contribuire all'evoluzione dei contenuti della rete realizzando podcast, caricando video personali, pubblicando foto, condividendo informazioni sotto forma testuale, e così via. Tra le attività di "produzione", però, ce n'è una che è una prerogativa specifica del Web 2.0: l'annotazione delle risorse della rete.

Anche in questo caso va precisato che le informazioni in rete vengono annotate da molto tempo prima dell'avvento del web 2.0, anzi, ad essere precisi, il principale meccanismo di annotazione dei contenuti, cioè l'HTML, è nato esattamente insieme al web. L'HTML, però, è un linguaggio di annotazione (o di markup) che si occupa di modellare semplicemente gli aspetti di formattazione dei contenuti indicizzati in rete, senza alcun accenno all'aspetto contenutistico e semantico.

L'impianto generale dell'HTML, inoltre, è di tipo spiccatamente sintattico. Nel set di base di annotazioni che l'HTML include non esiste – allo stato attuale – nessun tag finalizzato a modellare il significato dei contenuti inseriti all'interno della pagina.

Questo, fin dalle origini, è stato uno dei limiti principali del web, un limite che diventa sempre più importante al crescere delle informazioni presenti sulla rete. Proprio al fine di risolvere questo ed altri problemi simili, Tim Berners-Lee, l'inventore del web, aveva profetizzato una visione futuristica della rete in cui tutte le informazioni presenti su Internet fossero annotate in modo tale da associare a ciascuna di esse il corretto significato: il *Semantic Web*<sup>13</sup>.

---

<sup>13</sup> Tim Berners-Lee, *Semantic Web*, "Scientific American Magazine", 2001  
<http://www.scientificamerican.com/article.cfm?id=the-semantic-web>

Per una serie di motivi, purtroppo, la visione originaria di Tim Berners-Lee sta faticando ad affermarsi, nonostante gli sforzi iniziali. Individuare le motivazioni alla base di questo rallentamento non è semplice: la storia della rete (ma, più in generale, la storia di molte tecnologie) insegna che la definizione di alcuni standard non è direttamente collegata con l'utilizzo generalizzato degli stessi da parte degli utenti. Nel caso del web semantico la motivazione principale è legata proprio alla scarsa "utilità" percepita dagli utenti: allo stato attuale buona parte del popolo della rete non avrebbe grossi vantaggi dall'utilizzo, ad esempio, di formati come RDF per la modellazione delle informazioni.

Proprio alla luce di questo, erroneamente, inizialmente il web 2.0 era visto come una visione "semplificata" del web semantico, uno scenario in cui gli utenti si prendessero carico del compito di modellare le informazioni in modo tale da associare ad esse una qualche forma di annotazione semantica, ma – a differenza del web semantico - una semantica semplicemente comprensibile agli utenti, non alle macchine.

I meccanismi di annotazione delle informazioni nascono proprio in questo contesto, come uno strumento a disposizione degli utenti per associare agli oggetti presenti sulla rete (documenti, foto, bookmark, video, ecc.) una serie di termini specifici in grado di definirne in modo dettagliato contenuto e significato, anche senza usare una semantica formalmente definita come nel caso del web semantico. Nel gergo della rete questo tipo di attività viene ormai universalmente denotata col termine *tagging* e consiste appunto nell'attribuzione di una o più tag, finalizzate a modellare il significato, il contenuto o semplicemente delle informazioni caratteristiche sulla risorsa che si vuole annotare.



La pratica del tagging si sta quotidianamente affermando come un'attività sempre più diffusa per catalogare le risorse presenti in rete e proporre agli utenti altre informazioni correlate e di potenziale interesse.

Quasi tutte le applicazioni web 2.0 integrano ormai all'interno della propria piattaforma dei meccanismi di tagging, proprio a denotare il carattere trasversale di questa attività, svincolato dagli specifici domini applicativi.

Come inoltre dimostrato in numerose analisi, la quantità di risorse annotati in rete sta crescendo in maniera quasi esponenziale, ad un ritmo tale da rendere l'attività del tagging quasi un fenomeno sociale, da studiare ed analizzare al fine di comprendere l'impatto di questa attività sui classici meccanismi di organizzazione delle informazioni sulla rete.

In virtù del suo carattere trasversale, legato non solo all'informatica ma anche ad altre discipline slegate dall'ambito tecnologico, come ad esempio la semiotica e l'analisi del linguaggio, esistono allo stato attuale numerosi articoli che studiano il fenomeno del tagging, analizzandolo sotto diverse sfaccettature. In questa sede lasceremo da parte le analisi spiccatamente psicologiche, nonostante lo studio dei processi cognitivi e l'articolazione dei meccanismi mentali alla base dell'attività del tagging sia un argomento molto interessante ed ampiamente illustrato in altri articoli<sup>14</sup>.

## 2.2. **Dai tag alla folksonomy**

Il moltiplicarsi di sistemi di tagging sia nell'ambito della realizzazione di siti

---

<sup>14</sup> Rashmi Sinha, *A cognitive Analysis of Tagging*, articolo tratto dal blog ufficiale <http://rashmisinha.com/>

web, che nel loro utilizzo quotidiano da parte degli utenti ha portato alla diffusione del concetto di folksonomy.

Il termine è nato durante una discussione svolta il 24 luglio 2004 presso l'Information Architecture Institute fra Eric Scheid, Gene Smith, Thomas Vander Wal

Gene Smith, "architetto" di sistemi informatici, si chiese come si potesse chiamare il sistema di social classification basato sul tagging, diffuso in siti come Del.icio.us e Flickr. Eric Scheid rispose intuitivamente "folk classification" e Thomas Vander Wal si ispirò a questa risposta per il curioso neologismo "folksonomy", in cui, egli stesso specifica, il termine folk è da intendersi come "regular people", ovvero la massa di utenti medi del web.

Il passo decisivo per la diffusione nella blogosfera del nuovo termine fu il 3 agosto 2004, quando Gene Smith pubblicò nel suo blog personale un celebre intervento dal nome "Folksonomy: Social Classification"<sup>15</sup>, in cui descrisse l'episodio accaduto all'Information Architecture Institute e il suo parere sul fenomeno delle folksonomies.

"I think folksonomies can work well for certain kinds of information because they offer a small reward for using one of the popular categories (such as your photo appearing on a popular page). [...] On the other hand, I can see a few reasons why a folksonomy would be less than ideal in a lot of cases: none of the current implementations have synonym control (e.g. "selfportrait" and "me" are distinct Flickr tags, as are "mac" and "macintosh" on Del.icio.us), also, there's a certain lack of precision involved in using simple one-word tags--like which Lance are we talking about? And,

---

<sup>15</sup> [http://atomiq.org/archives/2004/08/folksonomy\\_social\\_classification.html](http://atomiq.org/archives/2004/08/folksonomy_social_classification.html)

of course, there's no heirarchy and the content types (bookmarks, photos) are fairly simple.”

Vander Wal si è poi preoccupato di chiarire la propria definizione di folksonomy come “il risultato del libero e personale tagging di informazioni e oggetti (qualsiasi cosa che abbia un URL) in modo da poterle poi facilmente recuperare”.

“The tagging – scrive Vander Wal - is done in a social environment (shared and open to others). The act of tagging is done by the person consuming the information.

The value in this external tagging is derived from people using their own vocabulary and adding explicit meaning (...). The people are not so much categorizing as providing a means to connect items and to provide their meaning in their own understanding.<sup>16</sup>

Il termine è stato quindi utilizzato dagli addetti ai lavori per indicare una tassonomia popolare o, come preferiscono dire in molti, una etnoclassificazione, cioè un sistema di categorizzazione collaborativa delle risorse sulla base di etichette (tag) assegnate direttamente dagli utenti finali, siano essi i produttori e/o i fruitori dei contenuti classificati.

Creare una folksonomy significa insomma mettere le etichette alle risorse per rendere più immediata la loro comprensione.

È una classificazione, o più precisamente, una indicizzazione, “dal basso”, creata dagli utilizzatori che attribuiscono un tag ad una risorsa messa sul web allo scopo di condividerla.

Quello della condivisione è un elemento importante nelle folksonomie; attribuire un tag tratto dal proprio vocabolario è necessario ma non

---

<sup>16</sup> Thomas Vanderwal 2005 - <http://www.dlib.org/dlib/january06/guy/01guy.html>

sufficiente perché si possa parlare di folksonomy. Secondo Vander Wal “il tag deve essere parlante” perché possa essere ricercato e recuperato da altri e permettere uno scambio fra culture e discipline diverse.

Fino ad oggi la creazione dei metadati per classificare le risorse bibliografiche e le risorse online è stata materia di professionisti, in quanto richiedeva l'applicazione di regole precise e dettagliate che trasformassero un linguaggio naturale in linguaggio “controllato” (si considerino gli standard internazionali come la ISO: 5964:1985).

Oggi sta emergendo un nuovo modo di classificazione e metadattazione, una classificazione “sociale” che non è più riservata ad una determinata categoria professionale; non c'è più bisogno di classificazioni pre-costruite, condivise in ambienti ristretti di professionisti catalogatori. Il navigatore che crea o segnala una risorsa sul web può attribuirgli dei “tag” liberamente. I tag non hanno bisogno di essere accreditati per esistere. Le informazioni non vengono gerarchizzate, ma raggruppate, aggregate.

Si tratta inoltre di una indicizzazione basata su parole chiave del tutto priva di struttura e di relazioni necessarie tra gli elementi; una catalogazione spontanea e collaborativa di risorse eterogenee che si adatta con facilità agli ambienti non gerarchici com'è la comunità del web.

Poiché non è verosimile che un'intera comunità di persone assimili in tempi ragionevoli un metodo comune di indicizzazione o classificazione delle risorse e lo applichi limitando gli errori, la folksonomy suggerisce che sfruttare le etichette spontanee poste da ciascuno sia comunque un notevole passo avanti rispetto al caos. Senza contare che la rete è un ambiente spesso privo di classificazioni precise, in cui elementi diversi potrebbero trovarsi nei contesti più impensabili.

Nella letteratura più recente si sono affermati rapidamente diversi approcci allo studio delle folksonomie e del social tagging. Alcuni autori hanno indirizzato i loro studi sull'attività del tagging, focalizzando la loro attenzione sul ruolo dell'utente nel mettere in relazione dei termini con delle risorse. Altri, interessati al vocabolario che ne viene fuori, hanno invece diretto i loro studi sulla folksonomia, intesa come insieme collettivo di termini assegnati da diversi utenti. Altri ancora si sono soffermati sull'aspetto sociale o tecnologico con cui le folksonomie sono costruite.

Si può quindi considerare

- il tagging come un processo (focalizzato sulla scelta dei termini da parte dell'utente)
- la folksonomia come il vocabolario collettivo che ne risulta (focalizzato sull'organizzazione della conoscenza)
- il social tagging come contesto sociale e tecnologico con cui il tagging ha luogo (focalizzato sulle reti e il social computing).

### **2.3. La fortuna del social tagging nel web 2.0**

La diffusione del social tagging è strettamente legata alla diffusione del web 2.0 e alla crescita esponenziale del cosiddetto 'user generated content': quando enormi quantità di contenuti informativi sono immessi in rete direttamente dagli utenti, gli utenti stessi tendono inevitabilmente ad essere anche i responsabili della descrizione dei contenuti immessi. Gli strumenti di caricamento di contenuti in rete devono permettere meccanismi semplici e funzionali di descrizione dei contenuti inviati e i tag sembrano essere la soluzione più semplice.

Il web 2.0 appare caratterizzato da quella "architettura della

partecipazione” teorizzata da Tim O’Reilly, secondo cui ad aggiungere valore ai servizi di rete sono gli utenti stessi, non necessariamente tramite strumenti espliciti, ma anche solo come effetto secondario dell’uso normale di un’applicazione.

La folksonomy è un sistema che esplode le sue potenzialità in particolari tipologie di archivi, dotati di caratteristiche quali una dimensione considerevole, la frequentazione da parte di una nutrita comunità di utenti e l’aggiornamento continuo.

In siti come YouTube o Flickr, che rispondono a queste caratteristiche, l’utilizzo di tassonomie più tradizionali non sarebbe in grado di rispondere tempestivamente alla magmaticità e alla mutevolezza dell’ambiente. Al contrario la folksonomy, pur non essendo in grado di offrire il miglior ordine possibile, permette il “miglior disordine possibile” garantendo una classificazione rapida, distribuita e scalabile, in grado di adeguarsi con precisione all’ambiente in continuo movimento (Maistrello, 2007).

La sua diffusione si deve soprattutto alle applicazioni internet basate sulle reti sociali, dentro cui i tag sono le parole chiave di qualunque condivisione; ai motori di ricerca specializzati, in cui essi operano come parole chiave volontarie e di nuova generazione; e ai blog, per i quali i tag sono diventati un’evoluzione delle preesistenti categorie.

Il successo della folksonomy merita attenzione e porta a fare riflessioni che vanno ben al di là del puro tecnicismo informatico<sup>17</sup>. In primo luogo, sorprende il successo di qualche cosa di inatteso, di non pianificato e di non sostenuto da alcuno degli attori forti del web. In un anno, o poco più, l’idea lanciata da Thomas Vander Wal di creare un sistema di tagging dal

---

<sup>17</sup> Nel dicembre 2005 il New York Time ha eletto la folksonomy tra le idee dell’anno - [www.nytimes.com/2005/12/11/magazine/11ideas1-21.html](http://www.nytimes.com/2005/12/11/magazine/11ideas1-21.html)

basso delle risorse in rete si è diffusa su scala planetaria dando vita a un insieme di siti che la adottano come base del proprio funzionamento e dei servizi che offrono.

I siti della folksonomy coprono aree diverse di esperienza e di conoscenza, hanno differenze anche significative tra di loro (dalla narrow alla broad folksonomy) e hanno conquistato il loro pubblico per le inedite possibilità di socializzazione di contenuti e di esperienze che offrono. Gli studiosi cominciano a discuterne dopo che il fenomeno si è ormai affermato e, anche se le idee alla sua base non sono nuove, la folksonomy non deriva naturalmente da esse, ma ne rappresenta la rielaborazione autonoma all'interno di quella comunità di attori del web che continua a essere la linfa del suo sviluppo.

Nel paesaggio digitale, la folksonomy porta con sé una vera rivoluzione che trova molte applicazioni soprattutto in ambienti in cui esiste già un linguaggio condiviso, come una banca dati di buone pratiche, una piattaforma e-learning, un settore di ricerca, un ambito professionale ma anche nella didattica.

Dicevamo che la folksonomy non è basata su idee nuove. Non è nuova infatti l'osservazione che nella loro vita gli esseri umani si scambiano non solo informazioni, ma anche criteri per classificarle e che anche in campi di conoscenza consolidati e maturi i sistemi di classificazione non sono fissi e univoci, ma evolvono continuamente riflettendo la diversità di conoscenze ed esperienze di chi opera in quel campo. La folksonomy implementa questo processo nel web, sfidando la potenza di calcolo e gli algoritmi dei motori di ricerca con un sistema a bassa efficienza ma capace di riflettere le dinamiche delle interazioni sociali. È vero anche,

però, che i siti della folksonomy sono, in generale, siti pubblici (l'accesso è libero e tutti sono invitati a visitare e apprezzare i loro servizi) che inducono il formarsi di comunità virtuali tra i loro utenti, non lasciandoli passivi ma offrendo loro, se li frequentano con continuità, spazi di protagonismo.

La socialità che la folksonomy attiva, quindi, in primis riflette l'accomunamento che nasce dall'essere utenti della stessa tecnologia, e solo al suo interno introduce distinzioni dipendenti dalla diversità di interessi e/o di esperienze. Per questo, il suo riferimento primo e principale è il web come spazio di esperienza ricco e completo, che essa concorre ad arricchire di servizi e possibilità di azione e interazione, ma trascura di misurarsi direttamente con le esperienze umane come si sviluppano nel tempo e nello spazio. Concludendo, la folksonomy è quindi un fenomeno interessante perché ha saputo rimettere in discussione il modello di organizzazione centralizzata delle risorse e delle loro sorgenti ma è ancora ben lontana dal costituire il nucleo di una nuova famiglia di applicazioni che espandono il ruolo della rete nelle interazioni sociali.

Rimane da chiedersi perché la folksonomy si sia affermata così velocemente. La rapida diffusione delle tassonomie popolari è probabilmente dovuta al loro approccio semplice, iterativo e distribuito all'organizzazione dell'informazione. In una parola, le folksonomies sono convenienti e in grado di riflettere accuratamente il modello mentale degli utenti, rispecchiando il loro linguaggio ed i loro bisogni informativi. Per questo la folksonomy si è estesa anche al di fuori di blog e social network, contagiando progressivamente il modo in cui sono pensati e strutturati gli altri contenuti del web. Già oggi stiamo assistendo ai primi esperimenti di utilizzo dei tag all'interno di musei reali, e le folksonomy iniziano a essere



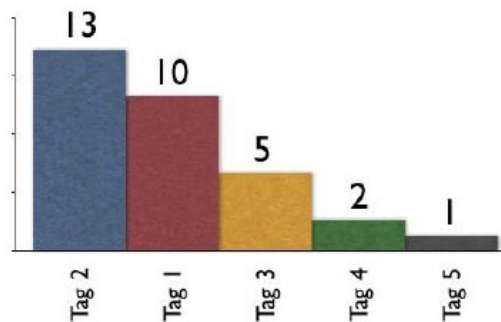
proposte da importanti testate online e da colossi quali Amazon come strumento complementare di accesso e organizzazione delle informazioni. In fondo, le folksonomy si limitano a copiare il funzionamento del cervello umano, e i possibili utilizzi sono limitati solo dalla nostra fantasia.

#### 2.4. **Broad folksonomy e narrow folksonomy**

I sistemi di classificazione semantica basati su folksonomies possono essere molto diversi l'uno dall'altro e le applicazioni oggi disponibili suggeriscono l'esistenza di due diversi approcci all'etnoclassificazione: una ristretta (*narrow folksonomy*) e una intesa in senso ampio (*broad folksonomy*). Nel primo caso, una persona mette le etichette su una risorsa; nell'altro, molte persone definiscono contemporaneamente la stessa risorsa.

Iniziamo a parlare della *broad folksonomy*, tecnica utilizzata da Delicious. La *broad folksonomy* è il risultato di molte persone che taggano lo stesso elemento e ogni utente può taggare l'oggetto in modo diverso, a seconda del proprio modello mentale, del proprio vocabolario e della propria lingua. In altri termini l'utente crea una risorsa e la rende accessibile agli altri. Gli altri utenti che sono interessati a questa risorsa le associano a loro volta tag corrispondenti ai termini che conoscono. Le persone trovano le informazioni sulla base dei tag.

La *broad folksonomy* permette di capire come viene taggato un oggetto in un ampio raggio.



Da questo grafico si nota che lo stesso oggetto è stato taggato con il tag “2” 13 volte, con il tag “1” 10 volte. Questo ci fornisce il trend dei tag per uno stesso oggetto etichettato. Inoltre è importante anche la fine di questa curva (*long tail*). In questa parte ci sono le persone che hanno taggato l’oggetto con un solo tag, ma comunque questo permette ad altre persone con mentalità simile di trovare l’oggetto, anche se questi tag sono poco utilizzati. (vedi paragrafo 3.5)

Una *narrow folksonomy* è invece il risultato di un minor numero di persone che taggano documenti digitali per facilitare un recupero o per la loro convenienza. Il tag utilizzato nel documento può essere sempre lo stesso, perché magari esiste solo quel termine per indicare quella risorsa. Quindi le *narrow folksonomies* hanno il vantaggio di essere più precise e facilmente recuperabile con una ricerca mirata.

La *narrow folksonomy* fa perdere la ricchezza della massa, non sfrutta l’effetto network, ma comunque fornisce benefici nel taggare documenti che non sono facilmente recuperabili con i normali strumenti di ricerca o taggati con tag ingannevoli. Flickr utilizza una *narrow folksonomy*.

La **narrow folksonomy** è diffusa soprattutto nei blog o negli archivi di immagini: ha il pregio di aprire alle risorse le porte dell’interazione con altri contenuti, moltiplicando le possibilità che una persona interessata li consulti. È l’autore, in genere, a definire esplicitamente i contesti a cui

ritiene che le proprie opere appartengano. Chi ne usufruisce può trarre giovamento dalla segnaletica messa a disposizione e dare avvio, a partire dalla risorsa etichettata, a nuove consultazioni mirate.

Ma è con la **broad folksonomy**, facilmente rintracciabile in social network come del.icio.us che questa pratica dà il meglio di sé. Ogni iscritto cataloga risorse identificabili in modo univoco (stringhe di indirizzi sul web, non necessariamente riferiti alle proprie creazioni): autori diversi possono mettere in circolo la medesima risorsa, definendola ciascuno secondo il proprio universo simbolico. Il fatto che lo stesso link sia segnalato da più persone contemporaneamente e risponda a tag diversi per ciascuna persona genera di per sé informazioni, permette alle risorse più interessanti di emergere e crea quel peculiare effetto moltiplicatore che caratterizza la broad folksonomy e favorisce più che mai la serendipity.

La differenza fra questi concetti è più chiara se prendiamo come esempi Flickr e del.icio.us, due siti che - come abbiamo visto - sono stati fra i primi a fondare il proprio valore aggiunto sui tag.

Flickr è un servizio di archiviazione e condivisione di fotografie. Ogni iscritto pubblica le sue immagini in un album personale consultabile da tutta la comunità. Fra le varie funzionalità disponibili, una delle più utili è la possibilità di associare a ciascuna fotografia alcuni tag, costruendo una sorta di indice analitico individuale di soggetti, luoghi o tecniche fotografiche. I tag di ciascuno confluiscono poi in un indice generale che moltiplica le possibilità che ciascuna immagine venga consultata da altri visitatori.

In altri termini se tutte le foto scattate a Parigi e inserite nella banca dati di Flickr riportassero, fra gli altri, anche il tag Parigi, chi cerca immagini che hanno a che fare con questa città non dovrebbe fare altro che selezionare

un link e scorrere un elenco di immagini a tema, senza il bisogno di scremare album per album tutte le fotografie presenti nel sistema.

La *narrow folksonomy* è pur sempre una forma di folksonomy, ma di una tipologia intermedia, perché una sola persona classifica un unico oggetto, inserendolo in un circuito di condivisione.

Si tratta di una pratica diffusa nel caso di risorse non testuali o difficilmente indicizzabili dagli algoritmi dei motori di ricerca perché favorisce il loro reperimento e la loro diffusione dividendo fra coloro che li immettono in rete il minimo sforzo di classificarli. Una fotografia non etichettata all'origine, infatti, diventa un costo per chiunque si trovi poi a doverla catalogare in un secondo momento, oppure una strada senza uscita in un ambiente che vive di connessioni fra le risorse.

Se invece si vuole salire di un livello nella condivisione di risorse si passa ad esempi di *broad folksonomy* come del.icio.us che, come dicevamo, è un servizio per il social bookmarking: ogni iscritto si appunta in uno spazio personale gli indirizzi di siti e documenti utili trovati durante la navigazione in rete e assegna a ciascuno di essi uno o più tag, creando il suo indice analitico di link personali. Anche in questo caso il sistema mette in comune tutti i collegamenti ipertestuali catalogati da tutti gli iscritti, moltiplicando gli effetti degli sforzi individuali.

Tante persone catalogano - a modo loro, con le descrizioni e le etichette che appartengono agli universi di senso di ciascuno - una stessa risorsa. L'indirizzo è una stringa univoca di testo: a un'identica sequenza di testo corrisponde una sola e ben determinata risorsa. La sua descrizione e l'assegnazione a una specifica combinazione di parole chiave contempla, invece, infinite variabili, che riflettono le specializzazioni, gli itinerari di senso e le interpretazioni personali che hanno portato a quell'indirizzo

ciascun utente.

Persone diverse che classificano una stessa risorsa, stabilendo nuovi agganci per connessioni di rete, fanno *broad folksonomy*, la pratica più vasta ed efficace di etnoclassificazione.

## 2.5. **Analisi dei tag: la coda lunga delle folksonomies**

Per valutare l'efficacia del social tagging sono stati condotti studi quantitativi sul comportamento degli utenti; tali studi sono applicabili alle cosiddette *broad folksonomies*, cioè quei servizi, come *del.icio.us*, in cui i tag assegnati a siti popolari possono essere anche centinaia.

Il primo risultato di rilievo di questa analisi è che la frequenza dei tag segue l'andamento tipico del principio di Pareto, la cosiddetta "legge 80/20", che è sintetizzabile nell'affermazione: la maggior parte degli effetti è dovuta ad un numero ristretto di cause

Applicando infatti questo stesso principio alla diffusione dei tag nei sistemi di social tagging si può affermare che poche parole sono molto utilizzate mentre la stragrande maggioranza lo è meno.

Questa legge sembra esprimere più in generale il comportamento degli utenti della rete sul web. L'intensa frequenza della parola "leader" indica un meccanismo di condivisione semantica che nasce anche dall'imitazione e dalla contrattazione implicita sui termini da adottare e che si realizza quando si visualizzano i tag utilizzati da altri utenti per designare la risorsa: è quindi presumibile che i termini statisticamente preferiti siano quelli ritenuti più idonei.

Secondo alcuni osservatori, la regola 80/20 è causata dai limiti del mondo

fisico, che tendendo a spostare l'attenzione dall'80% dei fattori meno influenti verso il 20% più influente, in qualche modo auto-sostenta questo rapporto.

La diffusione di internet e il superamento dei limiti dello spazio fisico ha però permesso di sperimentare nuovi modelli economici che mettono profondamente in discussione questo principio. Ad esempio la distribuzione del fatturato di amazon e di ebay è estremamente frammentata, ed entrambe le aziende raccolgono elevati margini proprio sugli articoli che sono distribuiti lungo la "coda lunga" del diagramma di Pareto.

L'espressione "coda lunga" (in inglese *long tail*) è stata coniata da Chris Anderson in un articolo dell'ottobre 2004 su "Wired" per descrivere proprio modelli economici e commerciali, come quello di Amazon.com o Netflix:

"What's really amazing about the Long Tail is the sheer size of it. Combine enough nonhits on the Long Tail and you've got a market bigger than the hits. Take books: the average Barnes & Noble carries 130,000 titles. Yet more than half of Amazon's book sales come from outside its top 130,000 titles. Consider the implication: if the Amazon statistics are any guide, the market for books that are not even sold in the average bookstore is larger than the market for those that are."<sup>18</sup>

La coda lunga è il fenomeno emergente nel mercato dominato da Internet, conseguenza del fatto che oggi gli utenti fruiscono (e acquistano) le risorse (musica, libri e altro) potendo disporre di un'offerta molto ampia dovuta proprio alla disponibilità delle risorse online e tramite Internet, che nella sua varietà va a soddisfare le esigenze di tutti, offrendo prodotti mainstream ma anche prodotti di nicchia.

---

<sup>18</sup> Chris Anderson, *The long tail*, "Wired", 12.10.2004  
<http://www.wired.com/wired/archive/12.10/tail.html>

Un cambiamento del tutto analogo sta avvenendo per quanto riguarda l'organizzazione delle risorse in rete e la "coda lunga" di Anderson porta dunque a ripensare la legge di Pareto anche per quello che riguarda le risorse in rete.

Più precisamente, la legge di Pareto rimane valida nelle proporzioni, ma cambia la distribuzione dei due segmenti della proporzione: l'80% del materiale, che rappresenta appunto la coda lunga, risulta poco consultato, ma mai "non consultato". Il punto è che non c'è motivo per non considerare anche questo 80%

Rossana Morriello nel suo articolo "La coda lunga delle biblioteche": spiega che nella coda lunga c'è di tutto: prodotti di qualità e di scarsa qualità, ma chi giudica cos'è di qualità? "Ciò che è di scarsa qualità per me, poiché non di mio interesse, può essere di interesse per altri"<sup>19</sup>.

Secondo la Morriello, dunque, non c'è più qualcuno che decide dall'alto, ma la selezione delle risorse in rete avviene dal basso e il filtraggio nelle nicchie serve a reperire ciò che è di interesse per me e altre persone che hanno il mio stesso interesse.

Nel mondo dell'informazione prima dell'avvento della coda lunga, la Morriello rileva quindi un approccio prescrittivo, in cui qualcuno decide a priori che cosa è di qualità, cercando di predire il gusto e spesso di crearlo e incanalarlo verso una certa direzione. Nel mondo a coda lunga, invece, gli strumenti informatici semplicemente misurano e descrivono il gusto, raccogliendo i feedback degli utenti, secondo un approccio di tipo descrittivo.

---

<sup>19</sup> Morriello, Rosanna, *La coda lunga delle biblioteche*, "Bollettino AIB", Vol. 49, n. 1, marzo 2009, pag. 69-82

Si tratta dello stesso tipo di rapporto che c'è tra tassonomie e folksonomie  
“Le prime, – spiega la Morriello – prescrittive, riescono a rappresentare la testa e il mainstream, le seconde, descrittive, riescono a rappresentare la coda per quanto frammentata e disomogenea sia.”<sup>20</sup>

Le folksonomie, dunque, partendo dal basso consentono di rappresentare le nicchie. Parafrasando si potrebbe dire che l'importanza del social tagging non si rivela solo in quei pochi tag molto usati, bensì proprio nella long tail di termini inattesi e non molto usati che costituiscono altrettanti accessi per un'ampia fetta di persone: si conferma così che la sovrabbondanza di accessi è una risorsa e non semplice “rumore”.

Da un'analisi più approfondita emergono altri modelli utili per comprendere le potenzialità del tagging. Il primo risultato indica che i tag utilizzati per primi hanno maggior frequenza, mentre i successivi tendono a decrescere: se si volesse dedurre un sistema di categorie condivise, bisognerebbe quindi considerare non solo i tag più frequenti ma anche quelli che sono stati usati per primi. Sembra qui confermarsi la valutazione secondo cui il comportamento degli utenti è caratterizzato dall'imitazione e dall'utilizzo di competenze altrui nella assegnazione dei tag.

Un ulteriore aspetto da sottolineare è che anche il contenuto dei tag, cioè i significati rappresentati da essi, sono stabili: si conferma così l'esistenza di una conoscenza condivisa e si spiega di conseguenza anche la rapidità di adeguamento, cioè la creazione di nuovi tag, quando compaiono concetti nuovi che prima non avevano nome. Una delle caratteristiche del linguaggio naturale consiste infatti proprio nella capacità di adattamento, e il social tagging si rivela infatti particolarmente capace di accogliere nuovi

---

<sup>20</sup> *ibidem*



concetti. Proprio questa maggior adattabilità al cambiamento ha spinto molti commentatori come Mathes e sottolineare la funzione del tagging come indicazione di “percorsi desiderati” dagli utenti, in contrapposizione alla astrattezza e rigidità dei sistemi formali delle tassonomie o dei vocabolari controllati.

## 2.6. **Dinamiche sociali**

L'analisi dell'uso che si fa in rete della folksonomy suggerisce riflessioni interessanti anche sulle dinamiche sociali che prendono vita nel web 2.0.

In primo luogo, all'interno dei meccanismi di funzionamento della folksonomy gli utenti del web 2.0 hanno la possibilità di ritagliarsi spazi di privacy, mettendo in campo pratiche di resistenza al controllo disciplinato dai fornitori dei servizi informatici. Basti pensare alle possibilità offerte dalla folksonomy per “targettizzare” i pubblici degli User Generated Content, utilizzando lo slang della sottocultura di una certa comunità per titolare e per assegnare tag a un video di YouTube.

Poniamo il caso, ad esempio, di pubblicare su YouTube un video che ha come protagonista il centravanti del Napoli Lavezzi, e di volerlo rivolgere ai tifosi di calcio, e in particolare alla comunità dei tifosi del Napoli. Una strategia sarebbe quella di applicare al video il tag “El Pocho”, che è il soprannome che è stato dato al calciatore nell'ambiente calcistico. Possiamo supporre che le persone che digitano nel motore di ricerca di YouTube “El Pocho” siano appassionati di calcio, e più probabilmente tifosi del Napoli.

La folksonomy offre inoltre la possibilità di fare aumentare il numero di visualizzazioni di un contenuto: se un utente di YouTube volesse diffondere più possibile un video di un argomento poco “cliccato” dagli utenti, potrebbe associare al suo video etichette più popolari, anche se non perfettamente coerenti con il contenuto (poi starà ai primi frammenti del video la capacità di suscitare la curiosità del pubblico). Per scoprire quali sono le etichette più popolari basta andare nella pagine dei video più visualizzati del giorno, e vedere quali sono i tag comuni dei video nelle prime posizioni (scopriremmo ad esempio che la domenica e il lunedì sarebbero tag calcistiche, altri giorni più legati a trasmissioni TV come il Grande Fratello, etc.). Infine, la folksonomy offre anche la possibilità di praticare il cosiddetto “pivot browsing”, ovvero la navigazione di un utente tra i tag di altri utenti ritenuti affidabili, come metodo per la ricerca di contenuti.

Le riflessioni sulla folksonomy, e in particolare quelle relative all'imprecisione della categorizzazione, contribuiscono però anche al ripensamento degli strumenti da utilizzare per indagare i nuovi ambienti digitali, e soprattutto negli archivi di contenuti audio-visuali.

Poniamo, ad esempio, che un ricercatore voglia effettuare una ricerca sulle rappresentazioni del conflitto israelo-palestinese su YouTube, rispondendo a domande come: quanti sono i video che trattano del conflitto israelo-palestinese? Quante volte sono stati visualizzati? Di quali argomenti specifici trattano? Da quali e quanti siti sono linkati? Quanti e che tipo di commenti hanno prodotto? Il problema a questo punto è: come recuperare una lista esaustiva di questi video, se la folksonomy rende tanto imprecise la classificazione e le ricerche dei contenuti? Digitando nel campo del motore di ricerca le parole chiave “conflitto israelo-palestinese”

appariranno alcuni video, ma sicuramente la lista non sarà nemmeno lontanamente esaustiva della popolazione sotto indagine. Pensiamo ad esempio alla guerra che ha avuto luogo in Palestina nel 1948: un ebreo sionista la chiamerà “guerra di liberazione nazionale”, mentre un palestinese la chiamerà “al-nakba”, la disfatta. E così via: molti degli episodi chiave di questa lunga guerra vengono nominati in modi differenti dai diversi protagonisti. Per risolvere questo problema e ottenere dei dati affidabili sull'intera popolazione di video sotto indagine, è necessario sviluppare tecniche di costruzione di un campione statisticamente rappresentativo. Una possibile tecnica per risolvere questo problema consiste nell'individuare, attraverso l'analisi dei testi ufficiali dei protagonisti del conflitto, le diverse modalità con cui ogni stakeholders nomina il conflitto e utilizzare queste parole chiave per avviare ricerche sul motore di ricerca del sito.

Per concludere, dietro agli apparenti tecnicismi dei criteri di archiviazione degli User Generated Content e del funzionamento dei motori di ricerca del web 2.0 si nascondono questioni rilevanti che riguardano gli usi “imprevisti” delle tecnologie che gli utenti utilizzano come risorse. Si nascondono questioni relative al potere, alle dimensioni della disciplina e del controllo del web 2.0

Basti pensare all'uso che è stato fatto di Twitter durante le manifestazioni di protesta in Iran, quando gli utenti, al ritmo di circa un post al secondo, hanno tenuto insieme le voci di quanti stavano seguendo quello che accadeva nella capitale grazie a tag come “Tehran” e “Iran election”.<sup>21</sup>

---

<sup>21</sup> Al fine di proteggere l'identità dei micro-bloggers che aggiornavano dall'Iran, il passaparola telematico ha suggerito che ognuno cambiasse nel proprio profilo personale la località di provenienza, segnando Iran e l'ora locale più 3.30 Gmt, in modo da mandare in tilt i sistemi di controllo.

In una video-intervista pubblicata sul sito di TED<sup>22</sup> Clay Shirky, professore alla New York University che da tempo studia gli effetti sociali ed economici della rete, spiega come Facebook, Twitter e altri social media aiutano i cittadini nei regimi repressivi a raccontare le reali notizie, scavalcando i censori e come la fine del controllo “top-down” delle notizie sta cambiando la natura della politica.

Molti studi hanno già messo in luce come, sia gruppi privati sia istituzioni pubbliche, cercano di controllare e monitorare l'accesso e gli usi di internet per scopi politici o economici. Lo dimostrano il recente sodalizio Obama / Google, che ha fruttato al neo presidente la miglior campagna elettorale mai dispiegata nel web 2.0, o le pratiche di marketing aziendale e politico che si rivolgono direttamente a questo ambiente (come l'*astroturfing*, ad esempio). Si nascondono infine questioni metodologiche, sulle quali il ricercatore si deve interrogare, soprattutto se utilizza i motori di ricerca del web 2.0 per recuperare i contenuti da analizzare, sia utilizzando tecniche standard che tecniche non standard.

---

<sup>22</sup>[http://www.ted.com/talks/clay\\_shirky\\_how\\_cellphones\\_twitter\\_facebook\\_can\\_make\\_history.html](http://www.ted.com/talks/clay_shirky_how_cellphones_twitter_facebook_can_make_history.html)

## CAP. III - FOLKSONOMY vs TASSONOMIE

### 3.1. Classificare per il web

Quando si progetta un sito web si utilizzano, a volte senza neanche esserne consapevoli, i principi della classificazione, così come farebbe un bibliotecario di fronte a una raccolta di libri. Ci si trova infatti a dover organizzare una grande quantità di informazioni (testi, immagini, audio, filmati, etc.), con lo scopo principale di consentirne un facile recupero. Nella maggioranza dei casi, la soluzione più efficace è quella di ricorrere al buon vecchio metodo della classificazione.

L'uso più evidente sul web dei principi della classificazione si può riscontrare negli indici sistematici, cioè in quei siti, detti anche *directory*, che fungono da cataloghi organizzati delle risorse online.

Fino a qualche anno fa servizi di questo genere erano forniti da Yahoo! o Virgilio, oggi ci sono invece directory aperte come dmoz o WWW Virtual Library; si tratta in ogni caso di indici sistematici che consentono il recupero dell'informazione attraverso la compilazione di uno schema di classificazione gerarchico-enumerativo e la successiva indicizzazione, sulla base di questo, delle risorse del web selezionate.

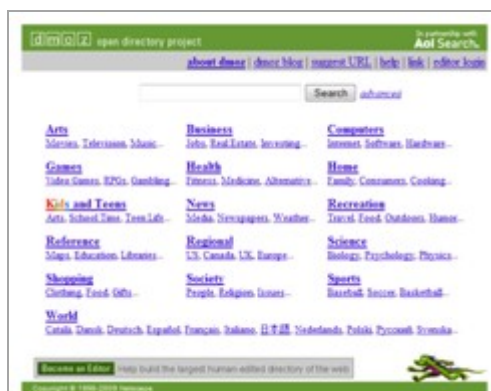
Compilare schemi enumerativo-gerarchici, selezionare le risorse e indicizzarle in maniera sistematica: niente di nuovo, quindi, rispetto al modo in cui la classificazione viene impiegata in biblioteconomia.

Generalmente tali indici sistematici hanno ripreso dalle classificazioni bibliografiche anche l'approccio semantico: i siti sono infatti raggruppati

prevalentemente a seconda dei loro contenuti.

Secondo Vittorio Marino, che ha affrontato il tema delle classificazioni per il web in un contributo per l'AIB<sup>23</sup>, il problema di un sistema realizzato in questo modo è che “la violazione della regola dell'unico *fundamentum divisionis* per ogni nodo della classificazione produce categorie non mutuamente esclusive e, quindi, uno schema organizzativo ibrido e incoerente, in cui più principi di divisione operano contemporaneamente.”

Un esempio di classificazione ibrida e incoerente si può ritrovare, ad esempio, nell'homepage di dmoz <www.dmoz.org>



Oltre alla categoria "Regional", spiega Marino, sono infatti presenti classi come "Kids and Teens" (classificazione per tipologia di utente), "World" (classificazione per lingua) o "Shopping" (classificazione per attività). Se un utente cerca un sito italiano che gli dia la possibilità di acquistare online un gioco per bambini, da quale categoria di primo livello deve iniziare? "Shopping", "Kids and Teens", "Games" o "Regional"? È probabile che siano previsti rinvii tra i vari rami della directory e che quindi si riesca ugualmente a arrivare alla categoria desiderata partendo da tutti e quattro i livelli possibili, ma l'imbarazzo causato da tale schema ibrido rimane.

<sup>23</sup> Vittorio Marino, *Classificazioni per il Web. I vantaggi dell'adozione di schemi a faccette*, AIB-WEB (Contributi), 2004 - <http://www.aib.it/aib/contr/marino1.htm>

Un'interessante ricerca sui principi di divisione nelle classificazioni del web è stata realizzata da Chaim Zins<sup>24</sup> nel 2002.

Lo studioso, dopo aver analizzato nove siti della rete (portali commerciali, directory, selezioni di risorse da parte di bibliotecari) con lo scopo di dedurre gli schemi organizzativi e il loro livello di "ibridazione", dimostra come "a mixture of categories that reflect different classificatory criteria" (Zins, 2002) sia la norma del web.

Ma sul web è necessario impiegare un solo principio di divisione alla volta? In altri termini, la mutua esclusività delle categorie e la loro coerenza sono proprietà irrinunciabili degli schemi di classificazione del web? Viene infatti da chiedersi se il rigore metodologico nella costruzione delle classificazioni, e quindi il rispetto delle regole e dei principi propri di un corretto procedimento classificatorio, sono requisiti fondamentali anche in rete o qui si possono mettere da parte a favore di un approccio più "disinvolto".

Una risposta a queste domande si può trovare nella teoria e nella prassi bibliotecarie. In biblioteconomia, infatti, il criterio dell'assoluta coerenza delle classificazioni è stato sempre subordinato al principio pratico secondo il quale è più importante che i sistemi di recupero dell'informazione siano utili agli utenti e facilmente usabili da questi piuttosto che scientificamente corretti.

Serrai<sup>25</sup>, ad esempio, sostiene: "nessuno ci assicura [...] che i metodi di suddivisione avvengano sempre sulla base di un unico *fundamentum*

---

<sup>24</sup> Chaim Zins, *Models for classifying Internet resources*, "Knowledge organization", Vol. 29, n. 1, 2002, pag. 20-28

<sup>25</sup> Alfredo Serrai, *Le classificazioni. Idee e materiali per una teoria e per una storia*, Firenze, Leo S. Olschki, 1977

*divisionis*, dato che nelle classificazioni bibliografiche sono state abbandonate le speranze di beneficiare della precisione e della assolutezza di una divisione logica"

Nonostante questo, nei sistemi bibliotecari rimane però la tendenza ad adoperare per ogni nodo della gerarchia un solo principio di divisione alla volta. Non è un caso che proprio le directory che utilizzano le classificazioni bibliografiche più diffuse come la Classificazione della Library of Congress, la Classificazione decimale Dewey o la Classificazione decimale universale, siano quei siti che, secondo i dati della ricerca di Zins, presentano un numero di principi di divisione minore. Nel web, invece, il non rispetto di questa regola fondamentale sembra essere costante.

Se sembra allora essere corretto, dal punto di vista dell'utente, rinunciare alle garanzie di una classificazione pienamente logica e scientifica, fino a che punto ciò è tollerabile? In altre parole, qual è il grado di incoerenza ammesso?

Per cercare una risposta a queste domande è necessario tornare alla prospettiva dell'utente per verificare se l'incoerenza degli schemi del web abbia realmente dei risvolti negativi sul recupero dell'informazione e possa essere così causa di ostacoli e difficoltà cognitive nella navigazione.

Secondo Zins questo è uno dei rischi che si corre:

"a structured classification scheme helps users to overcome the perplexing effect of the chaotic nature of the Internet by providing a cognitive model of information domain [...]. However, very often a cognitive model of the information domain can be misleading, since



it may be based on a biased, illogical, or inconsistent scheme".

Anche Rosenfeld e Morville<sup>26</sup> sono di parere simile:

"il potere di uno schema organizzativo puro deriva dalla sua capacità di suggerire un semplice modello mentale che l'utente può facilmente comprendere. Gli utenti riconoscono facilmente un'organizzazione specifica per audience o per argomento e schemi organizzativi puri abbastanza piccoli possono essere applicati a grandi quantità di contenuto senza sacrificarne l'integrità o sminuirne l'usabilità. Tuttavia, quando si inizia a miscelare elementi di molteplici schemi, ne consegue spesso confusione e le soluzioni sono raramente scalabili. Poiché le parti sono mischiate tra loro, non possiamo formarci un modello mentale"

Rosenfeld e Morville evitano però di assumere una posizione eccessivamente rigida e ammettono l'uso di schemi incoerenti in due circostanze: se il sito non classifica grosse quantità di contenuto o se le categorie create con *fundamenta divisionis* diversi sono fra loro divise.

Un modo possibile per seguire l'indicazione di Rosenfeld e Morville, e quindi separare le classi generate da criteri di divisione differenti, è proprio quello di realizzare un sistema basato su una classificazione "dal basso" che permetta di accedere alle informazioni tramite i tag.

### 3.2. L'accesso all'informazione tramite i tag

La diffusione degli strumenti di tagging collaborativo suscita una serie di interrogativi in termini di classificazione e indicizzazione: può una

---

<sup>26</sup> Louis Rosenfeld e Peter Morville, *Architettura dell'Informazione per il World Wide Web: Progettare siti Web Complessi*, O'Reilly - HOPS Libri, 2002

organizzazione della conoscenza costruita per fini personali diventare uno strumento di catalogazione efficace? Può funzionare una categorizzazione in linguaggio naturale, priva degli strumenti di controllo e di contestualizzazione dei vocabolari controllati e delle tassonomie create dalla Scienza dell'informazione?

A prima vista si sarebbe portati a dire di no. I tag vengono attribuiti da singoli utenti per organizzare le proprie risorse e questo specifico carattere individuale è rivelato dalla frequenza in rete di tag che indicano un compito come “*toread*” (da leggere), o che segnalano un giudizio di valore personale “*cute*” (interessante), ma anche nell'uso di espressioni che hanno significato solo per chi le scrive.

Gli strumenti di ricerca semantica consolidati hanno i loro punti di forza nell'organizzazione gerarchica e nella gestione delle relazioni semantiche tra le parole con il controllo delle varianti polisemiche, sinonimiche e grafiche.

La ricerca tramite la folksonomy è quindi priva di efficacia nella precisione e nel richiamo (*recall*) dei documenti pertinenti. Per sopperire a queste carenze sono state sperimentate varie tecniche come il raggruppamento, automatico o indotto manualmente, dei tag per definire “faccette” che in effetti realizzano servizi di ricerca efficaci ma solo in ambiti estremamente circoscritti e ben definiti. Se però dalla funzione del *find* si passa a quella del *discover*, il quadro cambia perché la navigazione nei tag correlati o nelle collezioni di utenti che utilizzano gli stessi tag, alla scoperta di relazioni semantiche orizzontali, cioè interdisciplinari, impreviste, si rivela molto efficace: questa ricerca che permette di fuoriuscire dagli ambiti di relazioni precoordinate prende il nome di *serendipity*.

Su questi temi si è svolto un dibattito abbastanza intenso. Molti hanno

sottolineato che la struttura gerarchica permette di contestualizzare disciplinarmente i termini fino a identificare nella catena, che va dal più generale al più specifico, la posizione esatta (*basic level*) del concetto che si vuole esprimere. Tuttavia le classificazioni gerarchiche hanno i loro problemi: perdono col tempo la corrispondenza con l'articolazione reale della conoscenza, richiedono perciò alti costi di manutenzione e un notevole sforzo per essere apprese e applicate.

In un articolo assai citato, dal significativo titolo *Ontology is overrated*<sup>27</sup>, Clay Shirky, professore alla New York University che si occupa da tempo di Internet e di condivisione di contenuti in rete, dopo aver sottolineato come limite delle tassonomie il legame con contesti storici, sociali e culturali definiti, che ne segnalano la provvisorietà e la parzialità, caratterizza le classificazioni bibliotecarie come finalizzate alla collocazione dei libri sugli scaffali per poterli ritrovare facilmente in base al contenuto. Ma poiché la collocazione del libro deve essere unica e univoca, il vincolo della fisicità sopravanza il suo contenuto intellettuale, costringendolo in una gabbia disciplinare rigida che lo rispecchia solo in parte. In realtà nelle classificazioni si fa confusione tra libro (contenitore) e concetti (contenuto), mentre Internet non ha vincoli fisici e il tagging rinvia ai significati.

Nei vocabolari controllati la gestione dei sinonimi prevede la normalizzazione di parole con significato simile verso un termine scelto. In realtà ogni termine sottende spesso significati diversi in contesti linguistici e sociali specifici, e la scelta di un termine rispetto a un altro comporta sempre in misura maggiore o minore una perdita di significato:

---

<sup>27</sup> Clay Shirky, *Ontology is overrated – Categories, links, and tags*, [http://www.shirky.com/writings/ontology\\_overrated.html](http://www.shirky.com/writings/ontology_overrated.html)

contrariamente ai vocabolari controllati, la logica del tagging non è binaria (sì/no) ma analogica, cioè prevede una parziale sovrapposizione di domini e quindi implica reali diversità di contenuto. La sinonimia costituisce quindi una ricchezza e non un difetto, e per quanto i vocabolari controllati sono *esclusivi*, il tagging è invece *inclusivo*.

In realtà, sottolinea Shirky, il “social tagging” è un’operazione di attribuzione alle risorse di “significati individuali” nell’ambito di “valori condivisi”. Perciò importante non è “chi categorizza meglio di me”, bensì “chi categorizza come me”, in definitiva chi ha valori e interessi simili ai miei.

Il fatto che il social tagging sia espressione della partecipazione degli utenti significa che il punto di vista rilevante nella valutazione di una risorsa non è solo quello che l’autore ha voluto dire – com’è di solito nella catalogazione mediata dal bibliotecario – ma quello che il lettore vede in essa, cioè il significato che le attribuisce: si pone cioè il problema del recepimento del contenuto del documento.

Tra gli estremi della relazione tra autore e lettore ci possono essere anche distanze considerevoli: il fatto che possano esserci interpretazioni giuste e significati sbagliati, se si escludono situazioni limite (estremo soggettivismo interpretativo e/o intenzionalità negativa), non toglie però nulla al fatto che nel “social tagging” si attua un processo di “sensemaking”, di attribuzione di significati in base alle preferenze e alle interpretazioni funzionali agli interessi e alle inclinazioni di una miriade di persone, e perciò per definizione non “false”, bensì “diverse”:

Tagging is fundamentally about sensemaking. Sensemaking is a process in which information is categorized and labelled and,

critically, through which meaning emerges<sup>28</sup>.

Rimane da valutare la critica di inaffidabilità delle scelte compiute dagli utenti dal punto di vista della qualità. In realtà l'approccio collaborativo all'organizzazione dell'informazione si rivela uno strumento efficace per limitare contenuti inaffidabili, promuovendo documenti "popolari" e certificati socialmente: in definitiva il social tagging rappresenta un modo per ricevere suggerimenti su argomenti di comune interesse, permettendo anche di identificare gli utenti più affidabili.

### 3.3. **Pro: l'accesso multidimensionale**

Ora che le folksonomy sono diventate così popolari, si sta iniziando a riflettere meglio sulle loro caratteristiche. Da un lato, il tagging può essere uno strumento molto utile per facilitare il reperimento delle risorse in rete, in un contesto in cui anche i migliori motori di ricerca non sono perfetti. D'altro canto, l'assegnazione di tag in modo del tutto arbitrario e spesso casuale non produce una vera classificazione. Lingue diverse usano parole diverse, per esempio, ma anche in una stessa lingua i termini e persino le idee con cui ciascuno definisce un argomento possono essere molto differenti. La maggior parte degli utenti, inoltre, non ha scrupoli e classifica anche il materiale relativo ad argomenti che non conosce affatto, rendendo caotica la situazione.

La condivisione di parole chiave mette a disposizione numerose modalità di consultazione. La prima, più ovvia: in ambito locale, ovvero all'interno di un singolo sito, l'insieme dei tag può dare vita a una mappa analitica dei contenuti attraverso la quale accedere ai post di un blog o alle fotografie

---

<sup>28</sup> *Ibidem*

contenute in un album in modo alternativo alla consultazione sequenziale consentita da archivi temporali e categorie tematiche.

Il secondo livello si genera invece con la raccolta di questi indici analitici a un livello più ampio, un gruppo di lavoro, gli iscritti a una community di blogging, i siti censiti all'interno di un motore di ricerca specializzato, e così via. Questo secondo approccio presenta gli stessi vantaggi, ma li traspone su una scala potenzialmente globale: tutti i contenuti prodotti da tutti i partecipanti entrano in relazione tra loro attraverso i legami che si creano tra parole chiave identiche. Inoltre, le relazioni semantiche che si generano in seguito alle combinazioni di diversi tag danno vita a grappoli di significati che selezionano in modo spontaneo gruppi di contenuti all'interno del web. La vastità delle interpretazioni possibili di una parola e la flessibilità del linguaggio connettono tra loro testi, immagini, brani audio o video che una classificazione tradizionale ad albero magari non avrebbe mai messo in relazione tra loro.

Ancora più interessante è il caso dell'interconnessione fra tag differenti. Nella maggior parte dei casi l'opera di classificazione spontanea dei contenuti porta all'inserimento di più di una parola chiave. Prendiamo ad esempio due post appartenenti a due blog diversi e non collegati fra loro: il primo post riporta come etichette i tag A, B e C, al secondo post sono invece assegnati i tag B, D e E. la condivisione del tag B mette in relazione le due unità di contenuto e, nello stesso tempo, amplia le possibili relazioni semantiche tra i significati A, C, D ed E.

Tutto ciò non a costi e in tempi così competitivi, poiché questi sono distribuiti tra tutti i partecipanti fino a diventare trascurabili. Sfruttando su vasta scala queste peculiarità, infine, le etichette permettono di tracciare

con rapidità l'emergenza di alcuni temi rispetto ad altri: dalle occorrenze di una certa parola all'interno di un gruppo di persone in un determinato arco di tempo, per esempio, si possono trarre implicazioni interessanti in fatto di argomenti caldi e idee condivise. È una sorta di attualità in cui le priorità sono determinate in modo democratico dal basso verso l'alto e non, come avviene nei mezzi di comunicazione tradizionali, dall'alto verso il basso.

In cosa differisce allora la folksonomy da altre forme di architettura dell'informazione come i sistemi a faccette o le tassonomie più formali? Tassonomie e sistemi a faccette sono progettati in modo organico e strutturato da professionisti del settore, cercando di intuire i bisogni degli utenti e le tipologie di contenuto prima ancora di iniziare a catalogare gli oggetti. In tal senso, esprimono la visione di un'autorità centralizzata e sono mirati a garantire precisione e sistematicità al corpus dell'informazione organizzata. Le folksonomy sono al contrario insiemi piatti di termini creati dagli stessi utenti nel momento in cui pubblicano, inseriscono e catalogano oggetti online. Non presuppongono la presenza di un'autorità centrale, ma, al contrario, lasciano emergere una visione decentralizzata e collaborativa.

Uno degli effetti più stimolanti del giocare con le analogie è la *serendipity*, ovvero il ritrovamento casuale di contenuti interessanti mentre si sta cercando tutt'altro, testimonianza spesso sorprendente di come mettere in collegamento tra loro i contenuti generi un valore spesso non preventivato. Benché gli algoritmi dei motori di ricerca siano sufficientemente potenti per indicizzare buona parte della produzione quotidiana del web, un sistema di classificazione come la folksonomy permette di recuperare anche informazioni residuali non altrimenti accessibili. A volte, per esempio, il tema principale di un intervento su un blog è talmente scontato da non

essere mai esplicitato in parole (come dimostrano ad esempio i commenti a caldo che seguono i cataclismi o i maggiori fatti di cronaca). Inoltre, un motore di ricerca tradizionale è in grado di trovare ricorrenze di parole, ma non ha la capacità di generare relazioni ipertestuali tra contenuti e tra persone.

#### 3.4. **Contro: polisemia e ambiguità**

La folksonomy però non è perfetta e, anche se viene migliorata giorno per giorno nella pratica e negli strumenti, continua ad avere dei limiti significativi, i principali dei quali sono: la polisemia, sinonimia, omonimia dei termini che possono essere utilizzati come tag dagli utenti; l'uso di idiomi differenti; la traslitterazione di parole o nomi da altri alfabeti; l'effettiva importanza e profondità dei tag assegnati dagli utenti.

Il limite più evidente è dato dall'ambiguità delle catalogazioni spontanee: persone differenti classificano in modo differente le stesse risorse.

Ciascuno di noi opera distinzioni di senso in funzione del proprio bagaglio culturale, della propria esperienza, della propria capacità di astrazione, del proprio modo di vedere il mondo. Per non parlare delle diverse lingue, dei sinonimi, dei luoghi comuni passeggeri, dei nomi che possono essere scritti in modi diversi pur essendo riferiti allo stesso concetto. Quella che dal punto di vista della classificazione può sembrare dispersione di energie e causa di entropia in un ambiente per sua natura già caotico, in realtà non fa altro che riprodurre dentro l'internet una struttura di organizzazione per analogie molto simile a quella che sta alla base del ragionamento umano.



Questi difetti hanno come principale conseguenza l'imprecisione della categorizzazione e, di conseguenza, della ricerca dei contenuti, soprattutto negli archivi in cui i tag assegnati dagli utenti hanno un peso maggiore nella classificazione.

Per fare un esempio, immaginiamo di voler cercare i post che trattano di filosofia all'interno di un blog: digiteremo la parola "filosofia" nel campo del motore di ricerca interno al sito, chiedendo al software di scandagliare, alla ricerca del termine richiesto, non solo i tag assegnati dall'autore, ma anche i titoli, i testi e i commenti. Il peso assegnato al "social tagging" è in questo caso limitato: anche se l'utente non etichetta il post con la parola chiave immessa nel motore di ricerca, se il contenuto stesso contiene la parola "filosofia", il post verrà listato e l'utente avrà la possibilità di trovarlo. Pensiamo invece di voler cercare i video di YouTube che parlano di sociologia: esattamente come nel caso del blog, immetteremo la parola "sociologia" nel campo del motore di ricerca interno al sito. Il software andrà alla ricerca del termine richiesto tra i tag, i titoli e i "testi informativi" (facoltativi), poiché il contenuto visuale e verbale dei video non viene scandagliato dal software. Il peso assegnato al "social tagging" è in questo caso evidentemente maggiore.

Il funzionamento dei motori di ricerca è molto più complesso della sommaria descrizione qui proposta, e i dettagli tecnici vengono abitualmente secretati dalle aziende che detengono il brevetto del software. I criteri secondo i quali un contenuto viene listato prima di altri, soprattutto in siti di natura commerciale, è una questione molto rilevante (da un punto di vista sociale, economico, politico), che riguarda la visibilità che un contenuto e il suo autore hanno rispetto agli altri. Tuttavia, anche senza entrare nei dettagli tecnici, è facile intuire il peso che la folksonomy ha nel funzionamento delle ricerche interne ai siti del web 2.0, e in

particolare quando i contenuti sono UGC audio-visuali.

Ricapitolando:

PRO	CONTRO
Serendipity: permette di trovare contenuto interessante anche senza una ricerca diretta di documenti rilevanti secondo una query	Ambiguità dei termini usati (Polisemia, sinonimia, uso del singolare/plurale, grafia, lingue)
Possibilità di visualizzare il dominio semantico dei tag	Assenza di contestualizzazione disciplinare
Abbassa le barriere della cooperazione: assenso su un significato non su una gerarchia	Nessuna uniformità di uso
Impulso a condividere	
Utilizzare i tag più popolari per identificare un vocabolario controllato	

### 3.5. Organizzare i tag per migliorare le folksonomies

Per limitare i problemi provocati dall'applicazione delle folksonomy nell'organizzazione delle risorse in rete, si è fatta rapidamente largo la convinzione che sono necessari modelli e strumenti per ri-organizzare i tag e mettere insieme la dinamicità dei tag liberamente scelti con la stabilità e la complessità dei vocabolari controllati.

Alcuni primi approcci sono stati già realizzati. Recentemente, le attività per pubblicare ed organizzare i tag sono state descritte come *tag gardening*, espressione introdotta per la prima volta dall'analista James Governor in un post del suo blog<sup>29</sup> e poi usata per descrivere i processi di

---

<sup>29</sup> Governor, James, *On the emergence of professional tag gardeners*, 10/01/2006 <http://www.redmonk.com/jgovernor/2006/01/10/on-the-emergence-of-professional-tag-gardeners>

manipolazione e ri-organizzazione delle folksonomies per renderle più produttive ed efficaci.

In particolare Isabella Peters e Katrin Weller, ricercatrici presso il Department of Information Science, dell'Università di Duesseldorf, hanno analizzato i diversi modi di revisionare e modificare i tag in base a diverse attività di "gardening"; fra cui modelli che permettono di aggiungere gradualmente strutture semantiche alle folksonomies e/o che le uniscono con le forme più complesse di sistemi di organizzazione della conoscenza.

"To discuss the different gardening activities, we first have to imagine a document-collection indexed with a folksonomy. This folksonomy now becomes our *garden*, each tag being a different *plant*. Currently, most folksonomy-gardens are rather savaged: different types of plants all grow wildly. Some receive high attention, others almost none. Some are useful for the community and retrieval tasks, others are not as they are highly personal or rather inappropriate for indexing purposes (e.g., the tag "me" or tags with spelling mistakes)."<sup>30</sup>

Uno dei primi metodi per rendere più facilmente accessibili e navigabili le folksonomies, secondo Peters e Weller, sono le *tag-cloud*, una forma di visualizzazione dei tag che mostra un gruppo di parole di grandezza differente. I tag più usati hanno un font più grande, quelli meno usati un font più piccolo. Lo scopo è quello di pesare la grandezza delle parole in base alla loro frequenza, in modo da visualizzare, a colpo d'occhio, quali sono gli argomenti più trattati. Oltre alla grandezza, le parole possono essere organizzate in ordine alfabetico, per similarità semantica, oppure

---

<sup>30</sup> Peters, Isabella e Weller, Katrin, *Tag Gardening for Folksonomy Enrichment and Maintenance*, "Webology", Vol. 5, n. 3, September 2008 - [www.webology.ir/2008/v5n3/a58.html](http://www.webology.ir/2008/v5n3/a58.html)

casualmente. Le voci sono di solito link che portano ad un elenco di oggetti collegati a quella voce.

Non si sa esattamente chi abbia inventato questo sistema. Secondo la versione inglese di Wikipedia il primo ad utilizzare le tag cloud è stato lo scrittore canadese Douglas Coupland nei cosiddetti “subconscious files” del suo romanzo *Microservi*, ma la pagina in questione non è più raggiungibile. Il loro utilizzo più importante è comunque quello fatto nei social media come Flickr, Del.icio.us o Technorati. Ma le tag cloud vengono ormai utilizzate ovunque.

Indubbiamente sono uno dei nuovi codici di comunicazione visuale che appartengono del Web 2.0, perché strettamente legate all'uso che gli utenti fanno dei tag. Ma sono in molti a chiedersi se funzionano davvero.

Sicuramente, come sostengono Peters e Weller, hanno il vantaggio di fornire all'utente, in maniera immediata, un elenco di qualificatori che sono stati attribuiti all'oggetto a cui si riferiscono. Al momento però non ci sono studi riconosciuti che attestino l'efficacia delle tag cloud come meccanismo di navigazione né tantomeno di architettura delle informazioni. Resta inoltre da capire quanto uno strumento del genere possa fungere da gadget alla moda per attrarre alcuni tipi di utenti, ma non si hanno molti dati a riguardo.

Una delle poche ricerche condotte su questo argomento è quella di Martin Halvey e Mark T. Keane<sup>31</sup>, due ricercatori della School of Computer Science and Informatics di Dublino che nel 2007 hanno presentato uno studio con sei diversi metodi per presentare elenchi di voci, fra cui le tag

---

<sup>31</sup> Martin Halvey e Mark T. Keane, *An Assessment of Tag Presentation Techniques*, Atti del 16th international conference on World Wide Web, maggio 08-12, 2007, Banff, Alberta, Canada

cloud. Gli intervistatori chiedevano agli utenti di trovare e selezionare alcune voci (città europee) in elenchi formati casualmente da una base di dati. Gli elenchi potevano assumere il seguente formato:

- nuvole di tag in ordine alfabetico
- nuvole di tag casualmente ordinati
- elenchi orizzontali alfabetici
- elenchi orizzontali casualmente ordinati
- elenchi verticali alfabetici
- elenchi verticali casualmente ordinati

Il risultato peggiore si è registrato proprio per le tag cloud. In realtà tutti i formati di presentazione che offrivano le voci disposte in maniera casuale, presentavano tempi mediamente più alti rispetto a quelli alfabetici. Ma anche fra gli alfabetici le tag cloud portavano a prestazioni peggiori.

Lo studio di Isabella Peters e Katrin Weller continua nell'analisi dei sistemi di tagging mettendo in luce un altro problema di base delle folksonomies: l'assenza di garanzia di ortografia e formattazione corrette dei tag inseriti dagli utenti.

“The very first activity in tag gardening would thus be *weeding*: tag weeding is the process of removing "bad tags" (e.g., tags with spelling mistakes). Elimination of spam tags should be the simplest form of tag weeding, and can probably even be performed automatically (to keep the image of gardening, this automatic spam removal could be characterized as using *pesticides*)”

Proprio come avviene nei giardini veri, però, l'identificazione dell'*erbaccia* non è sempre facile. Per esempio, si deve considerare quali tag possono essere rimossi da tutta la folksonomy (per esempio, le parolacce) e quali

dovrebbero essere rimosse soltanto da determinati documenti.

Le tecniche descritte dalle due studiose per “prendersi cura” dei tag e garantire così una migliore efficacia delle folksonomies sono diverse e complesse, arrivando a comprendere persino modelli che permettono di integrare gradualmente le folksonomies con delle strutture semantiche, unendole con le forme più complesse di sistemi di organizzazione di conoscenza.

Se dunque i primi studi sulle folksonomies mettevano l'accento sulla spontaneità e l'immediatezza del tagging collaborativo, col tempo la riflessione si è spostata decisamente verso un'analisi dei sistemi di organizzazione e strutturazione dei tag.

A questo punto pare dunque lecito chiedersi quale sia la reale differenza fra una tassonomia tradizionale e una tassonomia popolare.

### **3.6. La folksonomy è una tassonomia?**

L'idea che le folksonomies siano la nuova frontiera nel campo della classificazione dei contenuti è ormai assai diffusa fra osservatori e studiosi; ed è un'idea che si propone con forza a partire dal suo stesso nome, dato che - come abbiamo detto - la parola folksonomy è un neologismo composto da "folks" (popolare) e "taxonomy" (tassonomia) che sta ad indicare una classificazione "dal basso", creata dagli utilizzatori che attribuiscono tag alle risorse messa sul web al fine di condividerle. Le risorse non vengono quindi classificate a priori, ma aggregate dai navigatori/utenti.

Una volta chiarita la natura partecipativa e prettamente *bottom-up* del fenomeno, secondo Michele Santoro il termine che lo designa è viziato da un'ambiguità di fondo, perché "il significato di tassonomia (e di conseguenza quello di classificazione) è legato non alla presenza di descrittori in grado di esprimere un determinato contenuto, ma alla capacità di dar vita a distribuzioni categoriali di concetti"<sup>32</sup>.

Come infatti si legge nell'omonima voce di Wikipedia:

con il termine tassonomia (dal greco "taxinomia", dalle parole taxis = ordine e nomos = regole) ci si può riferire sia alla classificazione gerarchica di concetti, sia al principio stesso della classificazione. Praticamente tutti i concetti, gli oggetti animati e non, i luoghi e gli eventi possono essere classificati seguendo uno schema tassonomico.<sup>33</sup>

Santoro ritiene dunque che l'espressione "folksonomy" sia stata coniata in modo quanto meno affrettato, dal momento che gli utenti della rete danno vita a classi di concetti, ma individuano una serie di contenuti e ne danno una propria definizione, mentre il concetto di tassonomia (o, se si preferisce, di classificazione) implica sempre un'idea di distinzione, di ripartizione categoriale: quindi non una semplice descrizione di oggetti e fenomeni, ma un loro ordinamento in categorie più o meno omogenee.

In altre parole, il processo della classificazione dà vita a insiemi di concetti che vanno a raccogliersi in gruppi (le classi) a seconda del loro grado di somiglianza o di affinità, mentre separa quelli aventi caratteristiche diverse: tutti i membri di una classe devono infatti condividere almeno una caratteristica che i membri di altre classi non posseggono, e ciò permette

---

<sup>32</sup> Santoro, Michele, *Questa sera si cataloga a soggetto. Breve analisi delle folksonomies in prospettiva bibliotecaria*, "Bibliotime", anno X, n. 2, luglio 2007 <http://www2.spbo.unibo.it/bibliotime/num-x-2/santoro.htm>

<sup>33</sup> *Tassonomia*, in Wikipedia, l'enciclopedia libera - <http://it.wikipedia.org/wiki/Tassonomia>

di raggruppare oggetti o concetti fra loro simili distinguendoli da quelli diversi.

Tuttavia il processo di classificazione non si limita a raggruppare gli elementi simili e a separarli da quelli dissimili ma, all'interno dei singoli gruppi, permette di effettuare ulteriori ripartizioni, fino ad arrivare a elementi non più divisibili; si creano in tal modo delle "sotto-classi", all'interno delle quali si ritrovano caratteristiche che identificano in maniera via via più precisa i propri membri. L'atto del classificare allora non produce solo raggruppamenti più o meno omogenei di oggetti o di idee, ma una serie di suddivisioni di tipo gerarchico in base al grado di somiglianza o di affinità dei suoi componenti.

Niente di tutto questo avviene nel mondo delle folksonomies, dal momento che i tag vengono assegnati in maniera piatta e orizzontale, senza tener conto delle relazioni gerarchiche che intercorrono fra i concetti; ciò non consente di far emergere delle classi, e quindi non giustifica l'idea che, per quanto fatte dalla gente, si tratta comunque di tassonomie.

Ben più radicale - come abbiamo visto - è la posizione di Clay Shirky<sup>34</sup>, il quale sostiene che la categorizzazione degli utenti non rappresenta una estensione bensì una rottura radicale con le precedenti strategie di categorizzazione. Secondo Shirky infatti le classificazioni gerarchiche non rappresentano la realtà bensì sono costruzioni parziali e ideologiche, nate da vincoli fisici derivanti dal posizionamento dei libri sugli scaffali.

I tradizionali schemi di classificazione sono utilizzati quando le aree tematiche del contenuto da classificare sono identificate e limitate, e non cambiano nel tempo. Inoltre richiedono esperti catalogatori, una fonte

---

<sup>34</sup> Clay Shirky, *Ontology is Overrated: Categories, Links, and Tags*, 2005  
[http://www.shirky.com/writings/ontology\\_overrated.html](http://www.shirky.com/writings/ontology_overrated.html)



autorevole e utenti esperti nel catalogare.

L'utilizzo di un efficiente e completo sistema di classificazione richiede, da un lato, ai professionisti una comune visione chiara del dominio ed, dall'altro lato, utenti esperti per capire le categorie e la struttura della classificazione da utilizzare senza problemi.

D'altro canto, il contenuto su web, sempre più crescente, è formato da confuse fonti eterogenee di informazioni, mutevoli nel tempo, con un corpo non definito e non facile da classificare, destinato ad un gruppo variegato di persone, da soddisfare, con diversi background culturali e sociali in tutto il mondo.

È questo uno dei motivi che ha portato diversi appassionati sviluppatori di siti web e architetti dell'informazione a considerare la folksonomy come una scoperta rivoluzionaria, candidata a diventare un metodo di orientamento alternativo a quello dei motori di ricerca e che permette di arrivare alla conoscenza comune. Secondo questi "entusiasti" del social tagging, dato che il web è così vasto, non è possibile contare su esperti per avere il tempo di classificare tutto e spetta quindi agli utenti finali la possibilità di organizzare e trovare i contenuti che vogliono.

Si tratta indubbiamente di valutazioni eccessivamente ottimistiche e ancora premature, ma è certo che le folksonomy si stanno affermando come nuovo e interessante strumento di organizzazione delle conoscenze. Un sistema che viene generalmente identificato come una forma di classificazione, e per questo definito anche "etnoclassificazione" o "classificazione popolare", quando invece ad un'analisi più attenta sembra avere più punti in comune con la pratica della indicizzazione, e in particolare della indicizzazione per soggetto, piuttosto che con quella della classificazione.

Ciò che viene espresso tramite i tag è infatti l'argomento (o gli argomenti) di cui tratta la risorsa e non il campo del sapere a cui appartiene. Come l'indicizzazione per soggetto, la folksonomy fornisce accessi basati sul contenuto del documento e non su aspetti formali.

L'idea di base è che il contenuto di un insieme di risorse informative, disponibili sul web, venga indicizzato mediante i tag che gli stessi fruitori liberamente scelgono di assegnarvi. Un po' come se in biblioteca, restituendo un libro preso in prestito, il lettore consegnasse anche un'intestazione di soggetto che secondo lui ne rappresenta il contenuto, e questa venisse allora inserita nel catalogo.

Come l'indicizzazione, la folksonomy è soggettiva, dipende molto dalla sensibilità del catalogatore. Nel caso delle folksonomies, però, l'indicizzazione è affidata, oltre che agli autori, anche ai lettori stessi, ed è questa forse la reale novità derivante dalla pratica del tagging collaborativo. Tradizionalmente, infatti, l'indicizzatore è un terzo, un intermediario fra autore e lettore; e, come ha osservato Riccardo Ridi, la sua "terzietà" offre sia il valore aggiunto della competenza che quello della neutralità, almeno per quanto riguarda le motivazioni che lo muovono, allo stesso modo di un giudice che media tra accusa e difesa. Quando invece è l'autore a inserire i metadati, egli potrebbe alterarli a proprio vantaggio.

"La terzietà dell'indicizzatore rispetto ad autore e lettore non è solo una ottimizzazione per consentirgli di specializzarsi e di salvare il tempo del lettore, ma costituisce anche una garanzia che chi predispone i metadati abbia interesse solo a farlo nel modo tecnicamente migliore e non sia direttamente avvantaggiato - economicamente o da altri punti di vista - dal recupero di certi dati

piuttosto che di altri.”<sup>35</sup>

La questione di chi sia a compiere l'indicizzazione all'interno di un sistema di folksonomy assume allora anche un aspetto deontologico su cui ancora non si è riflettuto a sufficienza.

Riassumiamo qui alcune delle principali differenze tra folksonomy e taxonomy:

FOLKSONOMY	TASSONOMIA
Non ha una organizzazione gerarchica strutturata	I metadati sono strutturati (directory, path)
È un fenomeno sociale di tipo bottom-up che include il social bookmarking	Utilizza un sistema gerarchico o top-down
È una “user generated”, ovvero non c'è un ordine gerarchico dei termini che gli utenti devono seguire	È progettata da professionisti cercando di indovinare le esigenze degli utenti e tipologie dei contenuti
È possibile accedere ad un documento da più direzioni ed è possibile uscire dal percorso contestuale	Ha un'alta precisione, è finalizzata ad evitare ambiguità, in modo che la catalogazione sia controllabile
Non richiede manutenzione	Costosa da mantenere
Non richiede preparazione specifica	Difficile da applicare
Di immediata comprensione per l'utente	Difficile da capire per l'utente
Aggiornata in tempo reale	Rigida e non aggiornata tempestivamente
Post-coordinata	Pre-coordinata
Piatta	Multilivello
Contestualizza i concetti a livello di significati	Contestualizza i concetti tramite gerarchie disciplinari

A causa di queste differenze, le taxonomies e le folksonomies sono utilizzate in diversi potenziali campi di applicazione: le taxonomies sono adatte per una classificazione stabile, definita e limitata da una autorità

---

<sup>35</sup> Riccardo Ridi, *Metadata e metatag: l'indicizzatore a metà strada fra l'autore e il lettore*, relazione presentata al convegno “The digital library: challenges and solutions for the new millenium”, Bologna 1999  
disponibile su AIB-WEB <http://www.aib.it/aib/commiss/cnur/dltridi.htm>

centrale (classificazione Dewey) e da professionisti/utenti formati; le folksonomies invece sono utilizzate dalle persone comuni per motivi personali o sociali e sono vicine al linguaggio e al modello mentale degli utenti.

## **CAP. IV - ANALISI DI UN ESEMPIO: TRECCANI.IT**

### **4.1. Le enciclopedie in rete**

Fra la fine del secolo scorso e l'inizio di quello attuale la crescita del sapere ha raggiunto un ritmo esponenziale, determinando un continuo riassetamento della mappa delle discipline e del sapere. Di conseguenza le dimensioni, il taglio espositivo e gli ambiti disciplinari delle enciclopedie sono oggi estremamente articolati, come sono diversificate per età e per esigenze professionali sia le fasce del pubblico che esse si propongono di raggiungere sia l'estensione e il livello qualitativo delle informazioni che offrono.

Considerando il processo di parcellizzazione delle conoscenze, interno anche alle singole discipline più altamente specializzate, ci si chiede ora se le enciclopedie universali a stampa, che finora hanno avuto un primato indiscusso sulle varie scene nazionali, siano ormai soltanto i monumenti di un grande passato e se gli strumenti informatici non siano venuti a sopperire alla funzione delle enciclopedie di consultazione.

La struttura di un'enciclopedia ed il suo essere in naturale evoluzione sono infatti due proprietà particolarmente adatte a un formato digitale, da fruire sia su supporti di memorizzazione locale che in rete; di conseguenza tutte le maggiori enciclopedie a stampa, anche per motivi di concorrenza commerciale, si sono dovute progressivamente dotare di strumenti multimediali, fino a diventare enciclopedie digitali.

Tali pubblicazioni (basate prima su supporti CD-ROM e poi su DVD)

avevano il vantaggio di essere prodotte a basso costo ed essere facilmente aggiornabili; al contrario della forma stampata in genere includevano contenuti multimediali come animazioni, registrazioni audio e registrazioni video.

Ma le enciclopedie a stampa sono oggi esposte alla nuova concorrenza del web. Chiunque cerchi un'informazione rapida, o non abbia sotto mano un'enciclopedia aggiornata, cartacea o digitale che sia, sceglie sempre più spesso di andare a navigare in rete, per imbattersi in risposte più o meno complete e soddisfacenti.

Le enciclopedie consultabili in rete hanno una serie indiscussa di vantaggi rispetto alle opere a stampa, non ultimo quello di poter predisporre collegamenti ipertestuali tra voci concettualmente legate tra loro, cosa che permette di rendere assai più rapida la consultazione, o ancora quello di essere (potenzialmente) dinamiche: nuove informazioni possono essere mostrate quasi immediatamente, piuttosto che dover attendere la successiva pubblicazione su un supporto fisico.

Per fornire aggiornamenti tra le nuove edizioni numerose enciclopedie su carta pubblicano infatti dei supplementi annuali, come soluzione parziale al problema del mantenersi aggiornate, ma questo metodo richiede ovviamente al lettore lo sforzo aggiuntivo di verificare le voci sia sui volumi originali che sui supplementi annuali.

Inoltre l'informazione in un'enciclopedia stampata ha bisogno necessariamente di una qualche forma di struttura indicizzata. Tradizionalmente il metodo impiegato è quello di presentare le informazioni ordinate alfabeticamente secondo il titolo della voce.

Le enciclopedie in rete invece non sono legate alla necessità di imporre

una struttura predeterminata, perché è sufficiente che il lettore/utente digiti il termine di suo interesse perché il motore di ricerca gli restituisca i lemmi enciclopedici ad esso correlati.

Nonostante ciò però, la maggior parte delle enciclopedie in formato digitale offre un insieme di strategie di organizzazione delle voci, ad esempio per area di categorizzazione dell'oggetto o per ordine alfabetico.

Un'analisi più approfondita dei siti web delle principali enciclopedie disponibili in rete rivela la presenza di notevoli differenze all'interno del panorama internazionale fra le case editrici che mettono a disposizione online le proprie risorse e quelle che utilizzano il web esclusivamente per presentare la propria offerta commerciale.

In particolare per quanto riguarda la digitalizzazione delle risorse enciclopediche la scelta fatta nella maggior parte dei casi è quella di offrire brevi estratti delle risorse in forma gratuita e riservare invece esclusivamente agli utenti che sottoscrivono un abbonamento a pagamento la consultazione integrale dei materiali enciclopedici.

L'offerta di contenuti è particolarmente ricca e variegata in siti come quello della Enciclopedia Britannica o della Enciclopedia Encarta di Microsoft, che offrono articoli, spazi riservati a curiosità legate alla data, alla parola o al personaggio del giorno e dossier di tipo redazionale, oltre a materiali interattivi come video, audio e giochi online.

Le enciclopedie che hanno scelto di aprirsi a una reale forma di partecipazione dal basso e di collaborazione da parte degli utenti sono la francesca Larousse e la Enciclopedia Britannica, che prevedono una parziale forma di interazione con gli utenti e offrono una serie di strumenti dedicati alla comunità. Se infatti 'MSN Encarta' e 'Sapere.it' si limitano ad

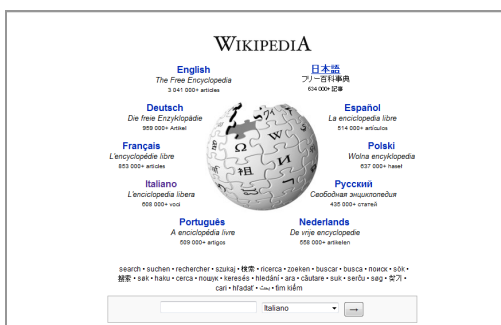
offrire la possibilità di segnalare un articolo a un amico, i due siti francese e inglese propongono agli utenti anche di contribuire e commentare i contenuti proposti e di interagire con i blog dei suoi autori commentando i post o partecipando ai forum.

In nessun caso, invece, è prevista l'associazione di tag alle voci enciclopediche, né da parte della redazione né tantomeno da parte degli utenti. L'unico esempio di enciclopedia in rete che presenti un sistema di folksonomy è offerto dalla Enciclopedia Italiana Treccani, tramite un sistema di tagging collaborativo che rappresenta un esperimento unico nel settore delle enciclopedie in rete.

Il sistema di tagging del portale Treccani sarà analizzato nel dettaglio nel prossimo paragrafo, mentre riportiamo qui di seguito una rapida rassegna dei siti delle principali enciclopedie presenti sul web.

## Wikipedia

<http://wikipedia.org>



Wikipedia rappresenta il più noto esperimento di enciclopedia online. Si tratta di una enciclopedia di tipo del tutto particolare, perché a contenuto libero e redatta in modo collaborativo da volontari e sostenuta dalla

Wikipedia Foundation, un'organizzazione no-profit.

A partire dal 2001 il sito ha raccolto quasi 4 milioni di voci in 200 lingue differenti, crescendo circa del 7% al mese e espandendo e contraendo in continuazione le sue voci.

Ogni voce dell'enciclopedia riporta una cronologia delle modifiche intervenute dal momento della creazione della pagina, ma tiene anche



traccia di tutte le discussioni che la redazione del singolo argomento ha provocato.

Anche il meccanismo di controllo dei contenuti è di tipo spontaneo e diffuso: chiunque può correggere eventuali inesattezze o intervenire sulle pagine esistenti, mentre la creazione di nuove voci è consentita solo agli autori registrati. Un gruppo ristretto di volontari, infine, è dotato di funzioni di amministrazione e può cancellare pagine o allontanare chi non rispetta il regolamento nelle discussioni.

Le voci dell'enciclopedia sono organizzate in una struttura gerarchico-enumerativa che prevede 5 grandi classi:

- Scienze matematiche, fisiche e naturali
- Arte · Letteratura · Arti visive e performative
- Scienze umane e sociali · Attività umane
- Società · Costume · Persone
- Tecnologia e scienze applicate.

È disponibile anche un indice alfabetico e, naturalmente, un potente motore di ricerca testuale, ma pur trattandosi di un esperimento “sociale”, non è prevista alcuna funzione di classificazione dal basso né sono presenti tag associati alle risorse.

Sulla affidabilità di Wikipedia e sui limiti di questo esperimento collaborativo hanno discusso in molti<sup>36</sup>.

Scriva Antonio Spadaro, una delle voci più autorevoli su questo argomento:

“Gli aspetti positivi e innovativi dei wiki diventano anche i limiti di questo sistema. La caratteristica essenziale del wiki, infatti, è

---

<sup>36</sup> Un elenco dei dibattiti più rilevanti è tenuto sulla stessa enciclopedia, nella voce “Wikipedia” [it/wikipedia.org/wiki/Wikipedia](http://it.wikipedia.org/wiki/Wikipedia)

l'apertura completa alla collaborazione, ma ciò significa, d'altra parte, che non esiste alcuna reale garanzia di validità e accuratezza dei contenuti immessi. [...] Questo ci fa capire che Wikipedia non può essere considerata come una tra le tante enciclopedie tradizionali [...] Ma soprattutto Wikipedia nasconde un'altra utopia, a suo modo ambigua: la democrazia assoluta del sapere e la collaborazione delle intelligenze molteplici che dà vita a una sorta di intelligenza collettiva. Questa utopia potrebbe nascondere una nuova forma di "torre di Babele", che ha il suo tallone d'Achille non solo nell'inaffidabilità, ma anche nel relativismo."<sup>37</sup>

Il paragone fra Wikipedia e le grandi enciclopedie cartacee che hanno costruito il loro prestigio sulla certezza del sapere rischia dunque di portare fuori strada la nostra riflessione, ma il successo di Wikipedia dimostra che un sistema basato sulle competenze delle persone e su una forma di collaborazione reticolare rappresenta un interessante strumento di raccolta e distribuzione della conoscenza.

---

<sup>37</sup> Antonio Spadaro, *Connessioni. Nuove forme della cultura ai tempi di Internet*, Bologna, Pardes Edizioni, 2006, pag. 101

## Brockhaus

<http://www.brockhaus.de>



L'enciclopedia tedesca Brockhaus Enzyklopaedie è presente in rete con un sito di tipo prettamente commerciale.

Nell'aprile del 2008, a causa della diminuzione delle vendite, si era deciso di realizzarne una versione online da consultare gratis e da

finanziare con annunci pubblicitari. In seguito, si è però rinunciato ad una pubblicazione in rete e nel febbraio del 2009 l'enciclopedia è stata assorbita dalla società Arvato/Wissenmedia GmbH, di proprietà della Bertelsmann.

Fino ad allora sul sito le risorse dell'enciclopedia erano disponibili online agli utenti che sottoscrivevano un abbonamento e i contenuti erano organizzati all'interno di un albero di navigazione a tre livelli. Era inoltre possibile fare ricerche attraverso un motore di ricerca per parola chiave all'interno delle diverse sezioni del sito.

Oggi invece, in seguito al brusco cambio di politica nei confronti del web da parte dell'azienda, il sito presenta esclusivamente l'offerta commerciale della casa editrice, senza la possibilità di consultare alcun tipo di contenuto enciclopedico né tantomeno alcuno strumento di partecipazione.

Un'occasione mancata forse, che testimonia il difficile rapporto delle "vecchie" enciclopedie nei confronti delle possibilità offerte dalla rete, auspiccate e temute allo stesso tempo.

## Enciclopedia Britannica

www.britannica.com



All'inizio del 2009 la storica Enciclopedia Britannica ha lanciato in rete una nuova versione del suo sito in chiave 2.0, spinto verso la partecipazione e la collaborazione.

Già da diverso tempo sul sito era possibile consultare le risorse

dell'enciclopedia; gli utenti possono accedere ai contenuti tramite una classificazione per soggetto che prevede 4 grandi classi di contenuti (History & Society; Arts & Entertainment; Geography & Travel; Science & Technology), ma anche tramite un indice o l'elenco alfabetico delle voci. È inoltre disponibile un motore di ricerca per parola chiave all'interno dell'enciclopedia o del dizionario.

A partire dal gennaio 2009 però gli utenti possono anche segnalare contenuti ed editarli direttamente sul sito, anche se prima della loro pubblicazione definitiva è necessario il placet della commissione di controllo, cui spetta l'insindacabile giudizio su quello che deve finire in pagina o meno.

Niente a che vedere quindi con il modello Wikipedia, dove invece è il lavoro collettivo di utenti esterni e redattori interni a confezionare le singole voci. Nel sito della Britannica infatti la sezione con i contributi degli utenti rimane separata e sono presenti tutti i nomi di chi ha contribuito. L'enciclopedia Britannica infatti non è mai stata a favore della politica della collaborazione democratica propria di Wikipedia, ritenuta troppo pericolosa per l'integrità e l'attendibilità delle voci.

Jorge Cauz, amministratore delegato dell'enciclopedia, ha dichiarato che

le modifiche che verranno effettuate dagli utenti saranno poi vagliate da un gruppo di editori e specialisti che si riserveranno di pubblicarle o meno. I migliori contributi potranno inoltre avere l'onore di essere pubblicati nella versione cartacea della nobile, antica e stimatissima Enciclopedia Britannica.

## Larousse

<http://www.larousse.fr>



L'editore francese Larousse ha reso disponibile nel 2008 una versione online della sua enciclopedia rendendo liberamente consultabili circa 150mila voci e 10mila immagini.

Allo stesso tempo è partita una sperimentazione in stile *wiki* di enciclopedia aperta a commenti e cointributi degli utenti, che però devono essere registrati e firmare i loro interventi.

L'iniziativa è ospitata in una sorta di sito parallelo a quello ufficiale; per ogni tema infatti si possono leggere a sinistra le voci di responsabilità diretta di Larousse consultabili gratuitamente e contraddistinte dal colore arancione; invece a destra i contributi dei lettori registrati, in viola. Né più né meno che articoli scritti e firmati dagli utenti.

Anche in questo caso siamo lontano dalla redazione in stile collaborativo di Wikipedia, ma la Larousse si è senza dubbio ispirata a un sistema wiki. La differenza sta nel fatto che i contributi del pubblico non sono totalmente aperti, o almeno non a chiunque, nè vanno a modificare e rimescolare quelli altrui: l'utente che vuole partecipare deve registrarsi e dichiarare

così la propria identità. Questo, come ha spiegato Line Karoubi<sup>38</sup>, direttrice del dipartimento "Dizionari ed enciclopedie" della Larousse, intervistata da *Le Monde*, garantirà non solo l'affidabilità e il rigore delle voci create, ma anche l'inaugurazione di nuovo modo per farsi notare. È infatti un segno di prestigio poter affiancare il proprio nome a una delle più importanti istituzioni linguistiche francesi.

La Larousse in rete, infine, non presenta nessuna organizzazione predeterminata delle voci enciclopediche, ma offre un motore di ricerca all'interno del quale l'utente può cercare il termine di suo interesse.

## Sapere

<http://www.sapere.it>



L'enciclopedia multimediale online del gruppo De Agostini si presenta come una webzine di cultura generalista con articoli e percorsi di approfondimento e gallerie multimediali, (foto, immagini a 360°, animazioni, audio, video).

Oltre all'enciclopedia generale, il sito contiene altre enciclopedie tematiche (Arte, Geografia, Storia, Medicina, Scienze ecc.), 9 dizionari linguistici, un atlante fisico e politico.

È disponibile una webzine, il *Giornale del sapere*, dedicata al panorama culturale italiano e internazionale, eventi (mostre, conferenze, convegni), scoperte scientifiche, innovazioni tecnologiche, correnti e protagonisti della

<sup>38</sup> Larousse s'ouvre au Net, *Le Monde*, 13/05/08

cultura.

Le risorse dell'enciclopedia sono disponibili online in versione integrale solo agli utenti abbonati. È possibile fare ricerche attraverso un motore di ricerca per parola chiave oppure navigare per canali tematici.

## Universalis

<http://www.universalis.fr>



Il sito della Encyclopédie Universalis francese è una webzine di cultura e risorse interattive per la formazione che corrisponde alla versione online dell'edizione cartacea in 28 volumi, con l'aggiunta di numerose risorse multimediali, collegamenti ipertestuali, aggiornamenti quotidiani e

settimanali, elenchi di siti web selezionati e assistenza per le ricerche su Internet.

Le risorse dell'enciclopedia sono disponibili online in versione integrale solo agli utenti abbonati, con la possibilità di sottoscrivere un abbonamento gratuito di prova della durata di 7 giorni.

È possibile fare ricerche attraverso un motore di ricerca per parola chiave, utilizzando anche funzionalità di ricerca avanzate, oppure navigare per canali tematici o sfogliando le risorse in ordine alfabetico.

## World Book Encyclopedia

<http://www.worldbook.com>



La World Book Encyclopedia è un'enciclopedia universale in lingua inglese in 22 volumi edita negli Stati Uniti dalla World Book Inc.

Dal 1998, oltre alla versione stampata e a quella in CD-ROM, la World Book Encyclopedia è presente in

rete con una versione online riservata agli abbonati chiamata *World Book Online*.

Il sito ad accesso libero è di tipo prettamente commerciale; presenta il catalogo e la rete di vendita, ma offre anche focus tematici, quiz e giochi interattivi, atlante geografico, suggerimenti di progetti e percorsi tematici per ragazzi.

Le risorse dell'enciclopedia sono digitalizzate e disponibili online in versione integrale solo agli utenti abbonati con offerte a pacchetti differenziati in base al target di riferimento: World Book Advanced, per studenti del college e delle scuole superiori, World Book Student, per studenti delle scuole inferiori, World Book Kids per bambini.



#### 4.2. La Treccani in rete: dalla classificazione per soggetto al social tagging

L'Enciclopedia Treccani è nata nel 1925 dalla collaborazione tra il filosofo Giovanni Gentile e l'imprenditore lombardo Giovanni Treccani per dare all'Italia, sull'esempio di altri paesi europei, un'enciclopedia nazionale, in grado di favorire la diffusione a tutti i livelli della cultura.

La Treccani ha rappresentato forse la massima impresa italiana di ricerca e per quasi un secolo l'istituto della Enciclopedia Italiana Treccani è stato un simbolo della cultura italiana. La sua opera principale, la *Enciclopedia Italiana di Scienze, Lettere ed Arti* è considerata un'opera prestigiosa e monumentale, che consta complessivamente di 54 volumi, per un totale di 55.000 pagine.

L'istituto della Enciclopedia Italiana Treccani è presente in rete dal 2001 all'indirizzo [www.treccani.it](http://www.treccani.it).



La prima versione del sito dell'istituto era focalizzata su un tipo di informazione culturale, dedicata in particolare ai temi della scuola e della lingua, ma con sezioni dedicate anche ad arte, cinema, diritto, storia, scienze e sport. I contenuti pubblicati erano tutti di tipo

redazionale o commerciale, di promozione del catalogo di opere dell'istituto, e non erano previste funzionalità di partecipazione da parte dei navigatori.

Nel 2008 l'Istituto dell'Enciclopedia Italiana ha deciso di rilanciare la sua immagine in rete pubblicando un nuovo portale e rendendo disponibili on-line parte delle sue risorse enciclopediche.

Oggi infatti gran parte delle risorse della Treccani sono online, consultabili gratuitamente. Sono disponibili circa 560mila lemmi, più o



meno lo stesso numero di Wikipedia Italia, aggiornati tempestivamente, se non in tempo reale, con accurate digressioni riguardanti argomenti di attualità.

Come nelle enciclopedie cartacee non mancano i riferimenti da una voce all'altra, anche nel portale Treccani l'ipertestualizzazione è pervasiva: ogni parola dei lemmi delle enciclopedie e del vocabolario è un rimando, ogni rimando è un approfondimento. Basta, infatti fare un doppio click su una qualsiasi parola del testo e si verrà subito rimandati alla relativa voce del Vocabolario online o della Enciclopedia online.

Sul sito è inoltre possibile commentare le voci enciclopediche, segnalare dei contributi, seguire percorsi tematici e ricerche incrociate, ma anche estendere la ricerca in rete in altri siti. Gli utenti possono partecipare alla comunità Treccani, un social network culturale dove creare un profilo personale, salvare percorsi di navigazione e link utili, scambiarsi messaggi, intervenire in dibattiti, ricevere aggiornamenti e porre quesiti

alla redazione. Il portale ospita inoltre una WebTv, con interviste a esperti su temi culturali e scientifici, video di divulgazione e approfondimento e convegni in streaming.

L'Istituto ha quindi deciso di mettere in rete un portale in pieno stile web 2.0 che risponde alla necessità di inserirsi in un contesto contemporaneo, grazie a servizi e strumenti caratterizzati da una diffusa partecipazione da parte degli utenti. Preso atto del fatto che la rete non rappresenta solo un ricco terreno di consultazione e indagine, ma anche un campo di azione vasto e multiforme, caratterizzato dall'interattività con i destinatari dei messaggi, l'istituto ha deciso di sperimentare nuove forme di comunicazione partecipativa, per fare del suo sito uno strumento di relazione, incontro e scambio delle conoscenze.

Digitalizzare e rendere effettivamente navigabile un patrimonio nato per la carta stampata - e cresciuto in diverse epoche storiche - è stata un'operazione complessa che ha richiesto un enorme lavoro di ottimizzazione per rendere agevole la fruizione via web di risorse nate e organizzate per opere a stampa.

Le voci enciclopediche disponibili in rete sono organizzate in una struttura gerarchico-enumerativa e suddivise in 5 classi principali:

- Scienze sociali e Storia
- Arte, Lingua e Letteratura
- Sport e Tempo libero
- Scienze naturali e matematiche
- Tecnologie e Scienze applicate

alle quali sono stati affiancati tre canali redazionali dedicati alla Scuola, alla Lingua Italiana e al Diritto.

Un'organizzazione dei contenuti di tipo esclusivamente gerarchico, però, non soddisfaceva i criteri di trovabilità delle informazioni, sia perché il tipo di classificazione utilizzata non risulta sempre di facile comprensione per l'utente medio, sia perché per raggiungere i lemmi enciclopedici (le “foglie” dell'albero di classificazione) seguendo un percorso di navigazione esclusivamente verticale e gerarchico è necessario arrivare al 5° livello di navigazione.

L'Istituto dell'Enciclopedia Italiana ha quindi deciso di aprirsi alla rete permettendo agli utenti di inserire tag utili a identificare tutti i contenuti del portale (lemmi enciclopedici, lemmi del vocabolario, articoli redazionali e video della webTV), descrivendo l'argomento di cui si parla o sottolineandone un aspetto specifico e rendendo così possibile la classificazione e la ricerca.

Così come accade nei sistemi che abbiamo definito di *broad folksonomy* come del.icio.us, anche nel portale Treccani una determinata risorsa può essere annotata sia dal suo “autore” (chi la indicizza per la prima volta all'interno del sito), sia, utilizzando tag differenti, da altri utenti, permettendo in questo modo di far emergere in modo omogeneo e relativamente immediato l'opinione degli utenti.

Per poter aggiungere dei tag gli utenti devono essere registrati.

La funzione “Aggiungi un tag” si trova alla fine del testo; una volta selezionata l'utente può scrivere il termine da associare a

Da Enciclopedie on line

TagT: Bernini Gian Lorenzo, Maderno Carlo, Roma, architettura, barocco

**Borromini (o Borromino; propr. Castèlli), Francesco**  
**Borromini** (o *Borromino*, propr. *Castèlli*), FRANCESCO. - Architetto (Bissone 1599 - Roma 1667), col suo antagonista G. L. Bernini è una delle due più originali e importanti figure dell'architettura del sec. 17° in Italia. Dopo un soggiorno a Milano, venne forse nel 1614 a Roma, dove fu impiegato come intagliatore e scalpellino nella fabbrica di S. Pietro. Fu protetto, e avviato alla professione di architetto, da C. Maderno, suo conterraneo e parente, dopo la morte del quale (1629) il B. seguì a lavorare in S. Pietro (specialmente nella realizzazione del baldacchino) e al palazzo Barberini (scala ellittica, disegno delle grandi finestre) alle dipendenze del Bernini, col quale si pose ben presto in aperto contrasto. L'attività autonoma del B. comincia (1634) con la costruzione del convento e della chiesa di S. Carlino alle Quattro Fontane (ma la facciata della chiesa, del 1667, è l'ultimo suo lavoro). Seguono una cappella in S. Lucia in Selci (1638-39) e le trasformazioni del Palazzo Spada e del Palazzo Falconieri (1640 circa); il palazzo Carpegna (oggi accademia di S. Luca; 1635-50); il convento e l'Oratorio dei Filippini (1637-50); la tomba Merlini in S. Maria Maggiore (1644 circa); il restauro di S. Giovanni in Laterano (1646-49); i lavori con G. e C. Rainaldi per il palazzo Pamphili (1645-50) e per S. Agnese in Piazza Navona (1653-57); la chiesa di S. Ivo alla Sapienza (...)

Leggi tutto  
(3409 battute)

Categorie

Biografie

☒ Aggiungi un tag

Contribuisci a catalogare le risorse del sapere treccani assegnando un tag a questo contenuto. Inserisci la parola a cui pensi e seleziona una delle voci che ti appariranno in elenco.

Scrivi la parola:

Artaserse  
Artaud, Antonin  
Art Bulletin, The  
arte

Chiudi

Aggiungi un'altra parola

Inserisci

quell'argomento. Mentre digita la parola il sistema restituisce un elenco di voci suggerite, fra cui l'utente deve scegliere e cliccare sul bottone 'Inserisci'.

Una volta inserito, il tag appare alla voce TagT, sopra il titolo della pagina.

Non si tratta dunque di un sistema di tagging "libero", perché l'utente non può inserire a sua scelta la parola che preferisce, come accade ad esempio in [del.icio.us](http://del.icio.us), ma deve scegliere il termine da associare alla risorsa all'interno di un thesaurus definito e circoscritto, che corrisponde al lemmario delle risorse disponibili sul portale.

Ciò permette di eliminare i problemi di ambiguità dei termini tipici delle folksonomie (polisemia, singolare/plurale, grafia, lingue), e allo stesso tempo garantisce che ad ogni tag corrisponda una voce enciclopedica.

Cliccando su ciascun tag infatti si lancia una ricerca che permette di consultare tutti i contenuti associati a quel termine: l'utente può quindi consultare la definizione del termine cliccato oppure navigare in percorsi tematici e correlazioni creati nel tempo dagli utenti della comunità.

Il sistema di tagging collaborativo dell'Istituto dell'Enciclopedia Italiana va dunque ad affiancarsi all'organizzazione gerarchica ed enumerativa dei lemmi previsti nelle classi dei 5 canali tematici, dando vita ad un sistema ibrido in cui la classificazione enumerativa è stabilita a priori e garantisce la possibilità di navigare l'albero dei contenuti in maniera sistematica, mentre i tag associati alle singole risorse permettono al navigatore di seguire percorsi trasversali e multidimensionali all'interno del patrimonio enciclopedico.

Grazie alla collaborazione degli utenti, le risorse Treccani vengono quindi

identificate secondo criteri che sono essi stessi a scegliere: una categorizzazione collaborativa, spontanea e “dal basso” delle informazioni, che tende a riflettere quello che gli utenti cercano e vogliono e che, come abbiamo visto, rappresenta un caso unico nel panorama delle enciclopedie in rete.

#### **4.3. L'uso dei tag in Treccani.it**

Dal luglio 2008 ad oggi<sup>39</sup> sul portale Treccani.it sono stati inseriti circa 80.000 tag.

Il ruolo principale dei tag nel sito Treccani può essere sintetizzato da questi 3 elementi:

- descrizione - essere efficaci come descrizione del contenuto digitale
- reperimento - aumentare possibilità e precisione nel reperimento di una risorsa
- associazione - offrire link ad altre voci enciclopediche o contenuti digitali pertinenti.

Non c'è dunque nessun intento “classificatorio”, dato che, come abbiamo detto, la struttura gerarchico-enumerativa è stabilita dall'alto nelle classi di argomento, quanto piuttosto una volontà di trattamento e indicizzazione delle risorse per renderle accessibili.

Un tag dovrebbe dunque rispondere alle esigenze di descrizione, reperimento e associazione che abbiamo descritto, ma nella folksonomy creata in questi anni dagli utenti del portale questi tre criteri non trovano sempre un effettivo riscontro.

---

<sup>39</sup> dati aggiornati al 30/11/2009

Le cause vanno ricercate da una parte in una serie di limitazioni del sistema di tagging adottato, e dall'altra in una pratica di associazione dei tag da parte degli utenti poco pertinente e, di conseguenza, poco efficace.

### Tag disponibili

Il primo problema che si pone, per una indicizzazione efficace, sono i tag a disposizione per l'indicizzazione.

Prendiamo il caso della voce "Alessandro Manzoni", che appare indicizzata in questo modo:

Da Enciclopedie on line **TagT:** Accademia della Crusca, Beccaria Cesare, Carducci Giosue, Goethe Johann Wolfgang von, Graf Arturo, Leopardi Giacomo, Milano, Monti Vincenzo, Napoleone Bonaparte, Parini Giuseppe, Rosmini Serbati Antonio, Sforza Carlo, Stael-Holstein Anne-Louise-Germaine baronessa di, Verdi Giuseppe, Verri Alessandro, Verri Pietro, barnabita ordine, giansenismo, letteratura

### **Manzóni <-3>, Alessandro**

I tag complessivamente sono 19, di cui ben 14 sono nomi di persona, 2 nomi di istituzioni e città, e solo 3 tag che indicano temi. Questo pone un problema di disequilibrio nell'indicizzazione: essa è dettagliata e precisa per quanto riguarda i personaggi citati nella voce, ma invece molto vaga nel descrivere temi e argomenti. I tre tag *tematici* sono "barnabita ordine, giansenismo, letteratura". Di questi solo due sono sufficientemente puntuali, letteratura è infatti molto generico.

Per equilibrare questo sbilanciamento e verificare i tag disponibili, è necessario porsi nella prospettiva di un utente e cercare di inserire alcuni tag suggeriti dalla lettura della voce enciclopedia o di specificare i tag esistenti.

Non sono però disponibili termini che possono venire in mente come "I

promessi sposi", "scrittore italiano", "letteratura italiana", "letteratura dell'ottocento", "romanzo storico", etc.

L'indicizzatore è dunque, di fatto, portato ad utilizzare termini generici come "letteratura" e non "letteratura dell'ottocento", "romanzo" e non "romanzo storico", "poesia" e non "poesia civile".

Se consideriamo, per fare un altro esempio, la voce "Semiotica", vediamo che non è possibile inserire tag a semiologi citati nella stessa definizione come Jurij Lotman, a cui tuttavia corrisponde una voce enciclopedica. Manca anche la possibilità di inserire tag come "semiologia", "semiologo", "semiosi", che sono presenti come voci nel vocabolario.

### **Tag generici**

Spesso i tag inseriti sono termini molto generici, che descrivono la materia o la classe di pertinenza della voce enciclopedica piuttosto che il suo contenuto specifico.

Il tag "filosofia", ad esempio, corrisponde a 47 schermate, ovvero circa 470 risultati. La quantità dei risultati non è un problema se nella prima pagina ci sono quelli più pertinenti (come avviene su Google), ma nel caso di Treccani sono presentati in ordine alfabetico. Il criterio alfabetico porta ad avere nella prima schermata personaggi e temi secondari come "Abellio, Raymond" o "ab aeterno" e rende difficile trovare i lemmi più rilevanti.

### **Tag incrociati**

L'utente è spesso portato ad associare due parole per creare un'espressione di senso compiuto che descriva la risorsa che vuole indicizzare, ma il sistema tratta ogni tag singolarmente.

Consideriamo ad esempio la voce "Accardo, Salvatore",



Da Enciclopedie on line

**TagT:** Bach Johann Sebastian, Beethoven Ludwig van, Genova, Ginevra, Guarneri, Italia, Napoli, Paganini Niccolò, Pertini Sandro, Siena, Stradivari Antonio, Torino, Vercelli, Vivaldi Antonio, cittadinanza, concerto, direttore d'orchestra, orchestra, teatro, violino

## **Accardo, Salvatore**

di GN

**Accardo, SALVATORE.** - Violinista e direttore d'orchestra italiano (n. Torino 1941), originario di Torre del Greco da cui ha ricevuto la cittadinanza onoraria, è considerato il più grande

Fra i tag è presente "Italia" o "italiano"; sono indubbiamente qualità e proprietà del violinista, questo è un criterio necessario ma non sufficiente: mettere insieme il tag "Italia" e "violino", infatti, non crea il tag "violinista italiano".

### **Tag non pertinenti**

Consideriamo la voce "D'Alema, Massimo".

Da Enciclopedie on line

**TagT:** Consiglio dei ministri, Gramsci Antonio, Pajetta Giancarlo, Partito comunista italiano, Scuola normale superiore, Sraffa Piero, Togliatti Palmiro, Unita L , giornalismo, politica, vela

## **D'Alèma, Massimo**

Tra i tag è presente "vela", nel lemma però in effetti non si parla in nessun punto di questo aspetto della vita privata di D'Alema; il tag è inserito a partire da informazioni a disposizione dell'utente e non pertinenti con quelle del lemma.

### **Possibili soluzioni**

Un soluzione efficace sarebbe quella di allargare lo spettro dei tag disponibili anche ai lemmi del vocabolario, ampliando quindi notevolmente il thesaurus dei termini a disposizione dell'utente.

Anche nei limiti delle parole chiave disponibili, però, si può comunque cercare di costruire una folksonomy più efficace “educando” gli utenti all'inserimento dei tag, sfruttando la loro appartenenza ad una comunità ristretta per indirizzarli a seguire una serie di norme per una pratica di tagging più consapevole, che segua queste semplici norme:

1) Non utilizzare tag troppo generali, ad esempio non mettere 'Italia' nelle voci di tutti gli scrittori italiani o di tutte le città: questo produrrebbe correlazioni molto estese e generiche, poco utili alle ricerche degli utenti. Invece "Promessi Sposi" nella voce Manzoni è un'ottima scelta. Scegliere invece tag il più possibile specifici, quindi "filosofia analitica" e non "filosofia", "film storico" e non "film", "giansenismo" e non "cattolicesimo" per evitare che a un tag corrisponda un elenco troppo lungo di risultati. Se il tag specifico non è presente, ma è presente solo quello generale, cercare un altro tipo di tag.

2) Inserire tag relativi ad ambiti differenti. Scegliere tag che appartengono a diverse categorie, per evitare un disequilibrio nell'indicizzazione.

Le categorie potrebbero essere le seguenti:

- Tema e argomenti principali - ad es. "borsa" in "broker"
- Temi specifici della voce - ad es. "polinomio" in "algebra" ; ed anche discipline specifiche e non generiche ad es. "logica matematica" in "Carnap" e non logica o matematica
- Descrizione del personaggio il più possibile precisa - ad es. direttore d'orchestra, se non è disponibile l'aggettivo si può utilizzare un sostantivo, ad es. "sceneggiatura" in Zavattini, Cesare visto che "sceneggiatore" non è presente fra i termini del thesaurus
- Corrente, periodo storico, luogo geografico - quando è

rilevante rispetto al contenuto. Ad es "Carrara", per via delle cave di marmo, in Buonarroti, Michelangelo, ma non "Italia" in Calvino, Italo, troppo generica.

- Soggetto-oggetto - ovvero per un autore l'opera, per un inventore l'invenzione. Ad es. nella voce Gramsci, Antonio si può inserire come tag "L'Unità"

3) Il tag crea collegamenti reciproci tra voci. Ricordare che il tag messo alla voce A crea un link verso un elenco di lemmi che condividono lo stesso tag B, C, D, E, ma è anche un link da uno qualsiasi di questi lemmi verso un elenco che contiene A. Il tag funziona quindi come una sorta di biglietto da visita del lemma. Se è normale che nell'elenco che corrisponde al tag "pietra preziosa" sia presente "diamante", è invece meno evidente la presenza di "diamante" nell'elenco del tag "Russia".

4) Inserire tag che permettono collegamenti utili e originali per l'utente. Tutti sanno che Italo Calvino è italiano, quindi non è utile inserire come tag "italia", molti meno che "J. L. Borges" è un suo riferimento.

5) Inserire tag che offrono approfondimenti rispetto al contenuto della voce, anche temi non direttamente presenti nella voce enciclopedica o presenti in altri termini. Quindi se si sta parlando di Energia elettrica, è interessante un tag con "Agenzia internazionale per l'energia atomica".

6) Scegliere tag che siano significativi presi da soli, non solo nel loro insieme: non sono infatti rivolti ad un motore di ricerca che li incrocia ma ad utenti che li cliccano e li scelgono singolarmente. Quindi meglio scegliere "storico, film" piuttosto che aggiungere al tag "film" il tag "storia".

7) Valorizzare le proprie conoscenze. Pensare a correlazioni più o meno note ed esplicite, magari non citate nella voce ma comunque corrette e oggettive: autore e opera, inventore e invenzione, maestro e discepolo.

Sembra lecito chiedersi, a questo punto, se i tag non fungano da veri e propri descrittori e se la pratica del tagging non coincida con quella di una reale indicizzazione.

Come l'indicizzatore esamina un documento, sceglie i concetti da indicizzare e li traduce nella terminologia controllata usata dal sistema di informazione in cui opera, allo stesso modo l'utente Treccani esamina una risorsa, sceglie i concetti da associare ad essa e li seleziona dal thesaurus dei termini disponibili.

L'uso dei tag nel portale Treccani in realtà non ha lo scopo di classificare le risorse né tantomeno quello di indicizzarle, e infatti non esiste nessun sistema di consultazione delle risorse enciclopediche in base ai tag, né è prevista alcuna delle attività di *tag gardening* descritte da Isabella Peters e Katrin Weller per pubblicare ed organizzare i tag (vedi par. 4.5). L'obiettivo dichiarato è infatti solo quello di creare correlazioni fra le risorse, permettendo all'utente una navigazione trasversale fra i lemmi enciclopedici.

Anche mettendo in pratica le indicazioni ora descritte per migliorare l'efficacia della pratica di tagging da parte degli utenti, non si riuscirebbe infatti a creare una folksonomy realmente efficace e in grado di reggere il confronto con la tassonomia tradizionale. Per fare ciò sarebbe necessario

un ripensamento radicale dell'architettura delle informazioni del portale, una ri-progettazione dell'infrastruttura tecnologica che permetta di gestire il sistema di tagging dei contenuti in maniera più articolata e flessibile, sul modello di quei "sistemi ibridi" che si stanno affermando in Rete per mettere a fattor comune la correttezza formale delle classificazioni gerarchiche e l'originalità di connessioni generate "dal basso"<sup>40</sup>.

All'interno del portale Treccani.it, quindi, la classificazione sociale mediante tagging collaborativo non rappresenta un'alternativa ai modelli di classificazione tradizionali messi a punto in ambito bibliotecario, ma piuttosto uno strumento complementare da utilizzare accanto a questi ultimi, per venire incontro all'ampio spettro di esigenze e modelli mentali che un sito con molti contenuti deve soddisfare.

Se infatti durante la navigazione ogni utente attua molteplici strategie di ricerca, allora il migliore sistema di *information retrieval* è quello che supporta il maggior numero e tipo di strategie.

Gli studi sui modelli comportamentali di ricerca dell'informazione da parte dei navigatori web<sup>41</sup> hanno dimostrato come l'utente, effettuando una ricerca, inizia magari con un termine generico oppure con un riferimento essenziale e poi procede esaminando una molteplicità di risorse. Ogni nuova informazione in cui incappa gli suggerisce altre idee, nuove direzioni e, di conseguenza, una diversa percezione della propria ricerca.

---

<sup>40</sup> Alcuni esempi di integrazione fra asse verticale della classificazione gerarchica e asse orizzontale dei tag sono, ad esempio, il sito di e-commerce Amazon <[www.amazon.com](http://www.amazon.com)>, il programma della Apple per riprodurre e organizzare file multimediali iTunes <<http://www.apple.com/itunes/>>, e il sistema di content management Drupal <<http://drupal.org/>>.

<sup>41</sup> Sull'argomento si vedano le pubblicazioni di Peter Pirolli presso il PARC – Palo Alto Research Center

A questo ambito si riferisce in particolare il modello di *berrypicking* (letteralmente: raccolta delle bacche) teorizzato nel 1989 da Marcia Bates<sup>42</sup>. Il termine *berrypicking* applicato al settore dell'*information retrieval* indica infatti una fondamentale modalità con cui l'uomo cerca le informazioni (in ogni ambiente, sia fisico sia digitale) modificando continuamente la propria strategia a seconda dei nuovi ritrovamenti.

Secondo la Bates il comportamento dell'utente nei confronti di una risorsa informativa è paragonabile a quello di un animale nei confronti del cibo. L'animale che cerca il cibo attiva i propri sensi per determinare quale strada è quella che condurrà alla maggior fonte di cibo; allo stesso modo l'utente in cerca di informazione deve "sensibilizzarsi" per determinare quale percorso all'interno di un sito lo condurrà più vicino all'informazione cercata.

Quando siamo alla ricerca di informazioni il nostro percorso devia continuamente verso nuove traiettorie proporzionalmente all'acquisizione di nuove consapevolezze o all'emergere di nuovi obiettivi.

Trovare un'informazione allora non necessariamente coincide col cercarla; l'utente può navigare senza una meta precisa (fase di *browsing*), oppure improntare una ricerca mirata verso qualcosa di ben delineato (fase di *searching*), e secondo la Bates questi due stati della ricerca possono sovrapporsi continuamente durante la navigazione in uno spazio.

Seguendo il modello di *berrypicking* della Bates e paragonando un sito ricco di contenuti come il portale Treccani a un supermercato, possiamo dire che l'utente entra con l'intento di comprare qualcosa (*searching*), ma

---

<sup>42</sup> Marcia Bates, *The design of browsing and berrypicking techniques for the online search interface*, Graduate School of Library and Information Science, University of California at Los Angeles, 1989  
<http://www.gseis.ucla.edu/faculty/bates/berrypicking.html>

se ha tempo nessuno gli vieta di fare un giro fra i reparti per vedere i prodotti in esposizione (*browsing*).

Se a supportare l'esperienza di navigazione del *searching* ci sono da una parte il motore di ricerca e dall'altra le classi tematiche proprie di un'organizzazione gerarchico-enumerativa delle informazioni, a stimolare l'attività del *browsing* ci sono i rinvii incrociati dei tag correlati alle singole risorse.

L'uso dei tag in Treccani.it non ha quindi alcun intento classificatorio (né del resto potrebbe averne, data l'assenza di un sistema di tagging dinamico e flessibile), ma mira a creare correlazioni inedite fra le risorse, permettendo all'utente una navigazione non lineare fra i contenuti disponibili e agevolando la cosiddetta *serendipità*, ossia il ritrovamento inaspettato di informazioni interessanti che non erano però oggetto della ricerca.

## CONCLUSIONI

Sulla reale utilità delle folksonomies e sul loro ruolo all'interno del Web 2.0 si è dibattuto a lungo<sup>43</sup>. Inizialmente, così come erroneamente si riteneva che il web 2.0 rappresentasse una versione "semplificata" del web semantico, le folksonomies sono state inquadrate come uno strumento alternativo alle tassonomie, più democratico e più semplice da implementare.

La realtà, invece, sembra diversa. Tassonomie e folksonomies rappresentano due approcci diversi e non necessariamente alternativi alla soluzione dello stesso problema, l'organizzazione e la strutturazione delle informazioni, ed hanno entrambe vantaggi e svantaggi, per cui la scelta di operare nell'uno o nell'altro senso è totalmente personale, soggettiva, ed in generale legata ai requisiti richiesti dal contesto in cui lo strumento dovrà operare.

Come abbiamo visto, infatti, le folksonomies hanno dei limiti importanti: quelli più ovvi sono legati alle caratteristiche dei linguaggi naturali già analizzati, per cui i tag, così come tutte le stringhe, possono essere caratterizzate da polisemia, sinonimia ed altre relazioni semantiche e sintattiche. Un altro limite, poi, è legato alla totale assenza di relazioni gerarchiche tra gli elementi; lo spazio generato dall'insieme di tutti i possibili tag è uno spazio piatto, senza relazioni. L'utente può annotare dei documenti con il tag "cane", ma questi non rientrerebbero nei risultati della ricerca dei documenti con il tag "animali", sebbene il primo documento sia chiaramente attinente il tema della ricerca.

Si tratta di limiti importanti, che denotano di certo una scarsa efficacia a

---

<sup>43</sup> Emanuele Quintarelli, *Folksonomies, Power to the People* - ISKO Italy-UniMIB meeting, Milano, 2005 - <http://www.iskoi.org/doc/folksonomies.htm>



livello strutturale e poca precisione nei meccanismi di ricerca.

Nonostante questo, però, ormai si fa largo uso delle folksonomies. Le motivazioni alla base di questo non si possono determinare con certezza, ma le dinamiche della rete ormai dimostrano come questo strumento sia utilizzato in modo trasversale in numerose piattaforme.

Di certo l'assenza di un vocabolario predefinito è un valore aggiunto importante di questi sistemi. Nei sistemi di tagging collaborativo l'utente può strutturare l'informazione sulla base del contenuto che egli stesso vede all'interno, semplificandone l'organizzazione ed il ritrovamento. Allo stesso modo una curva d'apprendimento totalmente nulla, assieme al forte senso di comunità che l'utilizzo di questi sistemi provoca, possono essere valutati come fattori importanti per spiegarne il successo. Ed infine, ultimo ma non meno importante, l'effetto della "coda lunga" di cui abbiamo parlato, la stessa *long tail* che Tim O'Reilly ha sottolineato come fattore decisivo per determinare il successo delle applicazioni Web 2.0: la possibilità di raggiungere anche "la coda" degli utenti del web, gli individui di nicchia, quelli che si discostano dalla massa e che magari cercano informazioni particolari.

Se poi la scarsa precisione nella ricerca nelle folksonomies è senza dubbio un limite importante, lo stesso non si può dire del *browsing* e della *serendipity* di cui abbiamo parlato a proposito dell'uso dei tag in Treccani.it, cioè la possibilità di incontrare elementi imprevisti, non voluti, ma potenzialmente di interesse. L'associazione di tag da parte degli utenti permette infatti di mettere in luce collegamenti fra le risorse a cui gli indicizzatori tradizionali non potrebbero dare evidenza, creando legami insospettati e insospettabili.

Si tratta di una forma di indicizzazione per soggetto del tutto originale, di

tipo spontaneo e collaborativo, che tende a riflettere il punto di vista degli utenti sulle informazioni disponibili in rete, permettendo di mettere in evidenza quello che gli utenti cercano e vogliono secondo criteri che sono essi stessi a scegliere.

Con la folksonomy siamo dunque di fronte a un sistema che pone al centro l'utente, che indicizza imponendo punti di vista decentralizzati, priorità e interessi imprevedibili, portando alle estreme conseguenze la dimensione collaborativa del web.

C'è da sottolineare che le folksonomies non sono la soluzione ad ogni moderno problema di organizzazione e strutturazione delle risorse in rete; sono più semplicemente uno strumento utile, innovativo ed a basso costo da utilizzare nelle giuste circostanze.

Senza dubbio i rischi di confusione e la difficoltà di gerarchizzare le informazioni possono costituire ostacoli significativi sulla via dell'affermazione della tassonomia dal basso. I progressi però sono già evidenti e molte opportunità ancora inesplorate.

Il fenomeno merita tutta l'attenzione.

## BIBLIOGRAFIA

Aurnhammer, Melanie, Hanappe, Peter e Steels, Luc, *Augmenting Navigation for Collaborative Tagging with Emergent Semantics*, International Semantic Web Conference, Athens, Georgia, USA, 2006  
<http://iswc2006.semanticweb.org/items/Aurnhammer2006ve.pdf>

Aurnhammer, Melanie, Hanappe, Peter e Steels, Luc, *Integrating Collaborative Tagging and Emergent Semantics for Image Retrieval*, Collaborative Web Tagging Workshop, Edinburgh, Scotland 2006  
[http://www.ibiblio.org/www\\_tagging/2006/17.pdf](http://www.ibiblio.org/www_tagging/2006/17.pdf)

Bar-Ilan, J., Shoham, S., Idan, A., Miller, Y., & Shachak, A., *Structured vs. unstructured tagging – A case study*. Collaborative Web Tagging Workshop, Edinburgh, Scotland 2006  
[http://www.ibiblio.org/www\\_tagging/2006/12.pdf](http://www.ibiblio.org/www_tagging/2006/12.pdf)

Bates, Marcia, *The design of browsing and berrypicking techniques for the online search interface*, Graduate School of Library and Information Science, University of California at Los Angeles, 1989  
<http://www.gseis.ucla.edu/faculty/bates/berrypicking.html>

Beckett, Dave, *Semantics Through the Tag*. XTech 2006: Building Web 2.0, Amsterdam, The Netherlands 2006  
<http://xtech06.usefulinc.com/schedule/paper/135>

Benvenuti, Nicola, *Social tagging e biblioteche. Implicazioni e suggestioni di una 'classificazione generata dagli utenti che emerge attraverso un consenso dal basso'*, "Biblioteche Oggi", Vol. 25, n. 3, aprile 2007, p. 35-42

Berendt, Bettina e Hanser, Christoph, *Tags are not metadata, but "just more content" - to some people*. Weblogs and Social Media, Boulder, Colorado, USA, 2007 - <http://www.icwsm.org/papers/2--Berendt-Hanser.pdf>

Birdsall, William F., *Web 2.0 as a Social Movement*, "Webology", Vol. 4, n. 2, 2007 - <http://www.webology.ir/2007/v4n2/a40.html>

Bogers, Tonie, Thoonen, Willem e van den Bosch, Antal, *Expertise classification: Collaborative classification vs. automatic extraction*, 17th ASIS&T SIG/CR Classification Research Workshop, Austin, Texas, USA,

2006 - <http://dlist.sir.arizona.edu/1792/01/bogers.pdf>

Campbell, D. Grant, *A phenomenological framework for the relationship between the Semantic Web and user-centered tagging systems*, 17th ASIS&T SIG/CR Classification Research Workshop, Austin, Texas, USA, 2006 - <http://dlist.sir.arizona.edu/1838/01/campbell.pdf>

Campbell, D. Grant e Fast, Karl V., *From Pace Layering to Resilience Theory: the Complex Implications of Tagging for Information Architecture*. IA Summit 2006, Vancouver, Canada  
[http://www.iasummit.org/2006/files/164\\_Presentation\\_Desc.pdf](http://www.iasummit.org/2006/files/164_Presentation_Desc.pdf)

Darin L. Stewart, *Building Enterprise Taxonomies*, Berkeley, New Riders, 2008

Farrell, S., Lau, T., & Nusser, S., *Building Communities with People-Tags. Human-Computer Interaction*, in "Human-Computer Interaction – INTERACT 2007", Springer Berlin / Heidelberg, 2007, pag. 357-360  
<http://www.springerlink.com.myaccess.library.utoronto.ca/content/e541852635m2561w/fulltext.pdf>

Feinberg, Melanie, *An examination of authority in social classification systems*, 17th ASIS&T SIG/CR Classification Research Workshop, Austin, Texas, USA, 2006 - <http://dlist.sir.arizona.edu/1783/01/feinberg.pdf>

Fiorentini, Barbara. *Il social bookmarking nel servizio di reference*, "Bibliotime", XI, 1, marzo 2008 <http://didattica.spbo.unibo.it/bibliotime/num-xi-1/fiorenti.htm>

Gnoli Claudio, Marino Vittorio, Rosati Luca, *Organizzare la conoscenza. Dalle biblioteche all'architettura dell'informazione per il Web*, Milano, Tecniche Nuove, 2006

Golder, Scott A. e Huberman, Bernardo A., *The Structure of Collaborative Tagging Systems*, "Journal of Information Science", Vol. 32, n.2, 2006, pp. 198-208 <http://www.hpl.hp.com/research/idl/papers/tags/tags.pdf>

Halvey, Martin e Keane, Mark T., *An Assessment of Tag Presentation Techniques*, 16th international conference on World Wide Web", Banff - Alberta, Canada, maggio 2007, <http://portal.acm.org/citation.cfm?id=1242826>

Hayman, S.. *Folksonomies and Tagging: New developments in social bookmarking*, Ark Group Conference: Developing and Improving

Classification Schemes, Sydney, 2007  
<http://www.educationau.edu.au/jahia/webdav/site/myjahiasite/shared/papers/arkhayman.pdf>

Heyman, Paul e Garcia-Molina, Hector, *Collaborative Creation of Communal Hierarchical Taxonomies in Social Tagging Systems*, InfoLab Technical Report 2006-10  
<http://dbpubs.stanford.edu/pub/showDoc.Fulltext?lang=en&doc=2006-10&format=pdf&compression=&name=2006-10.pdf>

John, Ajita e Seligmann, Doree, *Collaborative Tagging and Expertise in the Enterprise*, Collaborative Web Tagging Workshop, Edinburgh, Scotland 2006 - [http://www.ibiblio.org/www\\_tagging/2006/26.pdf](http://www.ibiblio.org/www_tagging/2006/26.pdf)

Kelkar, Shreeharsh, John, Ajita e Seligmann, Doree, *An Activity-based Perspective of Collaborative Tagging*. International Conference on Weblogs and Social Media, Boulder, Colorado, USA 2007  
<http://www.icwsm.org/papers/2--Kelkar-John-Seligmann.pdf>

Kipp, Margaret E., *Exploring the context of user, creator and intermediate tagging*, IA Summit 2006, Vancouver, Canada  
[http://www.iasummit.org/2006/files/109\\_Presentation\\_Desc.pdf](http://www.iasummit.org/2006/files/109_Presentation_Desc.pdf)

Kroski, Ellyssa, *The Hive Mind: Folksonomies and User-Based Tagging*, InfoTangle 2005 <http://infotangle.blogspot.com/2005/12/07/the-hive-mind-folksonomies-and-user-based-tagging>

Maistrello, Sergio, *La parte abitata della rete*, Milano, Tecniche Nuove, 2007

Mathes, Adam, *Folksonomies - Cooperative Classification and Communication through Shared Metadata*, 2004  
<http://www.adammathes.com/academic/computer-mediated-communication/folksonomies.html>

de Maurissens, Isabel, *Folksonomy: una classificazione sociale del web. Dal caos originario ai frutti della collaborazione*, "IR-Innovazione e Ricerca", 2006 - <http://eprints.rclis.org/7574/>

Metitieri, Fabio e Ridi, Riccardo, *Biblioteche in rete. Istruzioni per l'uso*, 3.ed., Roma-Bari, Laterza, 2005

Merholz, Peter, *Metadata for the Masses*, October 19, 2004

<http://www.adaptivepath.com/publications/essays/archives/000361.php>

Mika, Peter, *Ontologies are us: A unified model of social networks and semantics*, "Journal of Web Semantics", Vol. 5, n. 1, 2007, p.5-15  
<http://www.cs.vu.nl/~pmika/research/papers/ISWC-folksonomy.pdf>

Millen, David R e Feinberg, Jonathan, *Using Social Tagging to Improve Social Navigation*, Workshop on the Social Navigation and Community based Adaptation Technologies, Dublin, Ireland, 2006  
[http://www.sis.pitt.edu/~paws/SNC\\_BAT06/crc/millen.pdf](http://www.sis.pitt.edu/~paws/SNC_BAT06/crc/millen.pdf)

Morriello, Rosanna, *La coda lunga delle biblioteche*, "Bollettino AIB", Vol. 49, n. 1, marzo 2009, p. 69-82

Morville, P. *Ambient Findability: What we find changes who we become*, Cambridge, O'Reilly Media, 2005

Noruzi, Alireza, *Folksonomies: (Un)Controlled Vocabulary?*, "Knowledge Organization", Vol. 33, n 4, 2006, pag. 199-203

Peterson, Elaine, *Beneath the Metadata: Some Philosophical Problems with Folksonomy*, "D-Lib Magazine", Vol. 12, n. 11, November 2006  
<http://www.dlib.org/dlib/november06/peterson/11peterson.html>

Peters, Isabella e Weller, Katrin, *Tag Gardening for Folksonomy Enrichment and Maintenance*, "Webology", Vol. 5, n. 3, September 2008 - [www.webology.ir/2008/v5n3/a58.html](http://www.webology.ir/2008/v5n3/a58.html)

Quintarelli, Emanuele, *Folksonomies: power to the people*, ISKO Italy-UniMIB meeting, Milan, June 24, 2005  
<http://www.iskoi.org/doc/folksonomies.htm>

Quintarelli, Emanuele, Resmini, Andrea e Rosati, Luca, *FaceTag: Integrating Bottom-up and Top-down Classification in a Social Tagging System*. "ASIS&T Bulletin", June/July 2007  
[http://www.asis.org/Bulletin/Jun-07/quintarelli\\_et\\_al.html](http://www.asis.org/Bulletin/Jun-07/quintarelli_et_al.html)

Ridi, Riccardo, *La Biblioteca come ipertesto. Verso l'integrazione dei servizi e dei documenti*, Milano, Editrice Bibliografica, 2007.

Rosenfeld, Louis, Morville, Peter, *Architettura dell'Informazione per il World Wide Web: Progettare siti Web Complessi*, O'Reilly - HOPS Libri, 2002

Santoro, Michele, *Questa sera si cataloga a soggetto. Breve analisi delle folksonomies in prospettiva bibliotecaria*, "Bibliotime", anno X, n. 2, luglio 2007 - <http://www2.spbo.unibo.it/bibliotime/num-x-2/santoro.htm>

Schmitz, Patrick, *Inducing Ontology from Flickr Tags*, Collaborative Web Tagging Workshop, Edinburgh, Scotland 2006  
[http://www.ibiblio.org/www\\_tagging/2006/22.pdf](http://www.ibiblio.org/www_tagging/2006/22.pdf)

Smith, Gene, *Tagging: People-powered Metadata for the Social Web*, Indianapolis, New Riders, 2007

Smith, Tiffany, *Cataloging and You: Measuring the Efficacy of a Folksonomy for Subject Analysis*. 18th Workshop of the American Society for Information Science and Technology Special Interest Group in Classification Research, Milwaukee, Wisconsin, 2007  
<http://dlist.sir.arizona.edu/2061/01/Smith%5FUpdated.doc>

Spiteri, Louise, *Folksonomies, the Web and Search Engines*, "Webology", Vol. 5, n. 3, September, 2008  
<http://www.webology.ir/2008/v5n3/editorial17.html>

Sterling, Bruce, *Order Out of Chaos*, "Wired", January 31, 2008  
<http://www.wired.com/wired/archive/13.04/view.html?pg=4>

Trant, Jennifer, *Studying Social Tagging and Folksonomy: A Review and Framework*, "Journal of Digital Information", Vol. 10, n. 1, 2009  
<http://journals.tdl.org/jodi/article/view/269/278>

Terdiman, Daniel, *Folksonomies Tap People Power*, "Wired", 2005  
<http://www.wired.com/news/technology/1,66456-0.html>

van der Velden, Maja, *Organising Development Knowledge: Towards Situated Classification Work on the Web*, "Webology", Vol. 5, n. 3, September, 2008  
<http://www.webology.ir/2008/v5n3/a60.html>

Weinberger, David, *Everything Is Miscellaneous: The Power of the New Digital Disorder*, New York, Times Books, 2007

Wichowski, Alexis, *Survival of the fittest tag: Folksonomies, findability, and the evolution of information organization*, "First Monday", Vol. 14, n. 5 - 4 May 2009 -  
<http://firstmonday.org/htbin/cgiwrap/bin/ojs/index.php/fm/article/view/2447/>

2175